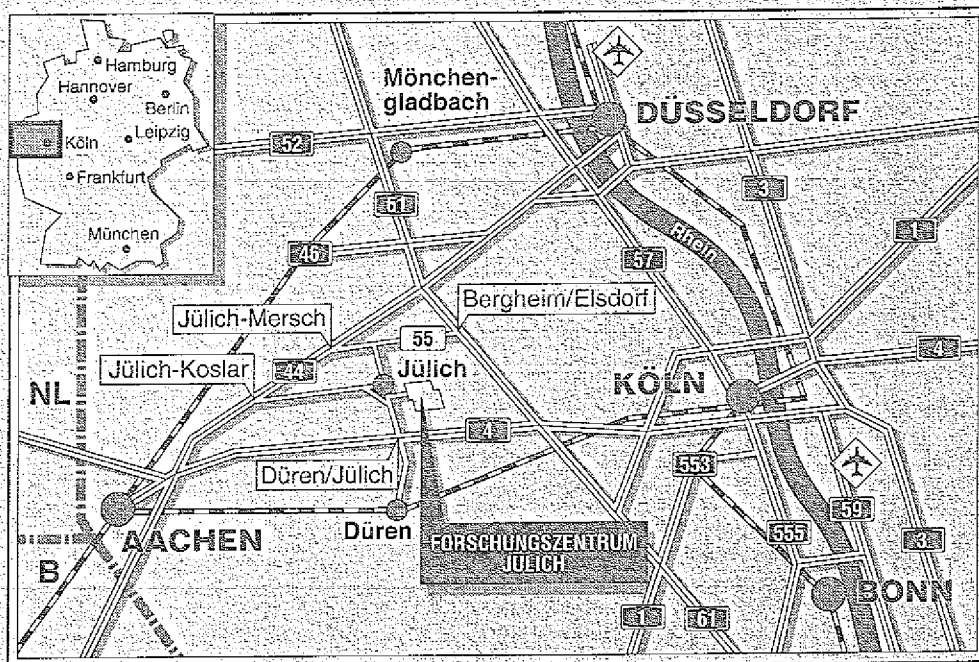


Institut für Festkörperforschung

**Methodische Untersuchungen zur
quantitativen Bildauswertung und zur
Rekonstruktion der Austrittswellen-
funktion in der hochauflösenden
Elektronenmikroskopie**

Andreas Thust



Berichte des Forschungszentrums Jülich ; 2978

ISSN 0944-2952

Institut für Festkörperforschung Jüli-2978

D82 (Diss. RWTH Aachen)

Zu beziehen durch: Forschungszentrum Jülich GmbH · Zentralbibliothek
D-52425 Jülich · Bundesrepublik Deutschland
Telefon: 02461/61-6102 · Telefax: 02461/61-6103 · Telex: 833556-70 kfa d

Methodische Untersuchungen zur quantitativen Bildauswertung und zur Rekonstruktion der Austrittswellen- funktion in der hochauflösenden Elektronenmikroskopie

Andreas Thust

[illegible]

Inhaltsverzeichnis

1	Einleitung	1
2	Modellstrukturen	6
2.1	Die Struktur von Ni_4Mo	6
2.2	Die Struktur von $\text{YBa}_2\text{Cu}_3\text{O}_7$	10
3	Grundlagen hochauflösender Abbildung	13
3.1	Aufbau eines Transmissionselektronenmikroskops	13
3.2	Abbildung kristalliner Objekte	16
3.2.1	Analyse der Objektstruktur	16
3.2.2	Synthese der Objektstruktur	25
3.3	Beschreibung durch numerische Bildsimulation	27
3.3.1	Simulation der Elektronenbeugung	27
3.3.2	Simulation des Abbildungsprozesses	27
4	Berechnung der Austrittswellenfunktion	28
4.1	Das Blochwellenverfahren	28
4.2	Das Multislice-Verfahren	35
4.3	Einsatz des EMS-Programmpakets	39
4.4	Ergebnisse	40
5	Kontrastübertragung	45
5.1	Kontrast des schwachen Phasenobjekts	45
5.1.1	Kontrast bei kohärenter Einstrahlung	45
5.1.2	Kontrast bei partiell kohärenter Einstrahlung	55
5.2	Nichtlineare Kontrastübertragung	68
6	Experiment	73
6.1	Präparation von Ni_4Mo Proben	73
6.2	Präparation von $\text{YBa}_2\text{Cu}_3\text{O}_7$ Querschnittsproben	74
6.3	Mikroskopie	75
7	Quantitativer Bildvergleich	80
7.1	Einführung	80
7.2	Theorie	82
7.2.1	Bildsimulation	82

7.2.2	Bilderkennung	83
7.2.3	Fourierfilterung	85
7.2.4	Symmetriekorrektur	87
7.3	Bilderfassung und Bildverarbeitung	91
7.4	Defokus- und Probendickenbestimmung	98
7.5	Kontrastübersicht	102
7.6	Diskussion	106
8	Rekonstruktion der Austrittswellenfunktion	110
8.1	Einführung	110
8.1.1	Zielsetzung	110
8.1.2	Bestehende Methoden	111
8.2	Theorie	114
8.2.1	Evaluation von Testwellenfunktionen	115
8.2.2	Der Simulated-Annealing-Algorithmus	116
8.2.3	Genetische Algorithmen	118
8.3	Rekonstruktion aus simulierten Abbildungen von Ni_4Mo	121
8.3.1	Rekonstruktion mittels des Simulated-Annealing-Algorithmus	122
8.3.2	Rekonstruktion mittels des genetischen Algorithmus	126
8.4	Rekonstruktion aus experimentellen Abbildungen von Ni_4Mo	129
8.5	Diskussion	131
8.6	Optimierung der numerischen Effizienz	136
8.7	Rekonstruktion aus Abbildungen von $\text{YBa}_2\text{Cu}_3\text{O}_7$	138
8.8	Diskussion	149
8.9	Ausblick	154
	Zusammenfassung	156
A	Fouriertransformation der Bildintensität	158
B	Explizite Darstellung des Transmissions-Kreuzkoeffizienten	160
	Literaturverzeichnis	161

Kapitel 1

Einleitung

Die Methode der hochauflösenden Transmissionselektronenmikroskopie erlaubt es, die Mikrostruktur fester Materie auf atomarer Skala zu untersuchen. Mit modernen hochauflösenden Mikroskopen ist es möglich, die lokale Atomanordnung innerhalb eines Festkörpers bis hinunter zu zwischenatomaren Abständen von weniger als $2 \text{ \AA}$ sichtbar zu machen. Durch Verbesserung der instrumentellen als auch der Auswertemethoden wird gegenwärtig ein Auflösungsvermögen von $1 \text{ \AA}$ angestrebt. Legt man realistische Elektronenenergien zugrunde, die noch eine zerstörungsfreie Beobachtung kristalliner Objekte ermöglichen und geht man von typischen elektronenmikroskopischen Präparaten aus, bei denen die vom Elektronenstrahl durchlaufene Strecke einige Nanometer beträgt, so ist eine Ortsauflösung von $1 \text{ \AA}$ allein aufgrund der quantenmechanischen Unschärferelation kaum noch zu steigern [1].

Im Gegensatz zu anderen Methoden der Mikrostrukturforschung wie Röntgenbeugung oder Neutronenstreuung ermöglicht die hochauflösende Elektronenmikroskopie eine direkte und lokale Beobachtung der atomaren Struktur eines Objekts in Form von Bildern. Hochaufgelöste elektronenmikroskopische Aufnahmen sind jedoch nicht mit lichtmikroskopischen Bildern vergleichbar und intuitiv oft schwierig zu interpretieren. Solche Aufnahmen gleichen vielmehr Interferogrammen, die die Information über atomare Objektdetails nur in verschlüsselter Form enthalten. Für die unerwünschte Verschlüsselung der Objektinformation sind zwei Prozesse verantwortlich, die im folgenden geschildert werden:

Zum einen unterliegt der Durchgang des einfallenden Elektronenstrahls durch das zu untersuchende Objekt quantenmechanischen Gesetzmäßigkeiten. Der einfallende Elektronenstrahl kann grob durch eine ebene Welle beschrieben werden (Siehe Abb. 1.1). Diese ebene Welle wird durch die Wechselwirkung der Elektronen mit den elektrostatischen Potentialen der Atome im Objekt auf komplizierte Art und Weise modifiziert. Dieser Vorgang wird als dynamische Elektronenbeugung bezeichnet. Die resultierende Wellenfunktion an der Austrittsfläche des durchstrahlten Objekts ist eine im mathematischen Sinne komplexe Funktion deren Amplitude und Phase die Information über die atomare Konfiguration des Objekts enthalten. Weder Amplitude noch Phase entsprechen hierbei notwendigerweise einer einfach verständlichen Projektion der Objektstruktur in Richtung des einfallenden Elektronenstrahls.

Die zweite Ursache für die schwierige Interpretierbarkeit hochaufgelöster Aufnah-

men liegt darin begründet, daß die Wellenfunktion an der Austrittsfläche des Objekts nicht direkt beobachtet werden kann. Ähnlich zur Lichtoptik wird das Wellenfeld in der sogenannten Bildebene mittels Elektronenlinsen fokussiert und abgebildet. Hierbei geht ein Teil der Information über die Objektstruktur verloren, da statt der komplexen Austrittswellenfunktion, die durch zwei Werte, nämlich Amplitude *und* Phase charakterisiert ist, nunmehr in der Bildebene „nur“ ein Wert, nämlich die Bildintensität, gemessen werden kann. Zusätzlich wird das Wellenfeld beim Transfer von der Austrittsfläche des Objekts zur Bildebene durch technisch nicht zu vermeidende Unzulänglichkeiten des elektronenoptischen Abbildungssystems drastisch modifiziert und in seinem Informationsgehalt reduziert.

Im Gegensatz zur Lichtoptik gibt es hierbei keine ideale Fokussierung, die ein „scharfes“ Bild liefert. Die variable Fokussierungsmöglichkeit elektronenoptischer Linsen bewirkt, daß von ein und derselben Austrittswellenfunktion eine Vielzahl gänzlich unterschiedlicher Abbildungen erzeugt werden können. Jede dieser Abbildungen beinhaltet einen anderen Teil der ursprünglich in der Austrittswellenfunktion vorhandenen Information und keine von ihnen kann als die „richtige“ angesehen werden.

Um dennoch verlässliche Aussagen über die Objektstruktur machen zu können, bedient man sich der numerischen Bildsimulation. Hierbei werden von einem Strukturmodell ausgehend die zuvor erwähnten kontrasterzeugenden Prozesse, nämlich die dynamische Elektronenbeugung und die elektronenoptische Abbildung, mittels eines Computers „nachgebildet“ (Abb. 1.1). Das eingehende Strukturmodell beinhaltet die auftretenden Atomsorten sowie die Koordinaten der einzelnen Atome. Durch Vergleich der auf der Basis des Strukturmodells errechneten Bilder mit den experimentell aufgenommenen Bildern kann die Richtigkeit des Modells überprüft werden. Gegebenenfalls muß das Modell geändert werden und der Vorgang so oft wiederholt werden bis eine zufriedenstellende Übereinstimmung zwischen experimentellen und simulierten Bildern besteht.

Der Vergleich zwischen experimentellen und simulierten Bildern wird üblicherweise durch den Betrachter „per Augenschein“ vorgenommen. Ein solcher „semi-quantitativer“ Bildvergleich [2] ist unkompliziert und für die Klärung einer bestimmten Klasse von Fragestellungen ausreichend. Problemstellungen, bei denen aufgrund physikalischer Vorbedingungen oder aufgrund von bereits erworbenem Vorwissen nur einige wenige Strukturmodelle mit deutlich unterschiedlichem Bildkontrast in Frage kommen, können auf diese Art und Weise geklärt werden.

Nicht mehr ausreichend ist der Bildvergleich per Augenschein bei Problemstellungen, bei denen es nicht auf die Entscheidung zwischen einigen wenigen wohldefinierten Strukturmodellen ankommt, sondern Größen ermittelt werden sollen, die kontinuierlich variieren können. Die Dicke des durchstrahlten Objektbereichs sowie die Linsenfokussierung sind Beispiele für solche variable Größen, deren genaue Werte aus dem Experiment oft nur unzureichend bekannt sind. Die gesuchten Werte werden meist dadurch bestimmt, daß man bekannte und periodisch aufgebaute Bereiche aus experimentellen Aufnahmen als Basis für ein Strukturmodell wählt und mit diesem Modell eine große Anzahl von Bildern berechnet, die sich durch die Parameter Fokussierung und Objektdicke unterscheiden. Hat man ein simuliertes

Bild gefunden, das mit dem experimentellen Bildbereich übereinstimmt, geht man davon aus, daß die zu dieser Simulation benutzten Werte für Objektdicke und Fokussierung den experimentellen Verhältnissen entsprechen. Die genaue Kenntnis der so ermittelten Abbildungsparameter erlaubt es, danach die materialwissenschaftlich eigentlich interessante Klärung unbekannter Strukturdetails als entkoppeltes Problem zu behandeln.

Aus einem Bildvergleich „per Augenschein“ resultieren in diesem Falle konkrete Zahlenwerte. Es ist klar, daß ein solches Verfahren aufgrund der Subjektivität des Betrachters unbefriedigend ist. Zudem ist die zu überprüfende Anzahl simulierter Bilder so groß, daß die Gefahr besteht, beim „Durchprobieren“ der möglichen Parameterkombinationen die beste Übereinstimmung zu übersehen. Quantitative Vergleiche zwischen simulierten und experimentellen Bildern, die diese Nachteile beheben, sind in der Literatur eine große Rarität [3, 4]. Ein halbautomatisches und äußerst schnelles Verfahren zur Bestimmung von Fokussierung und Objektdicke wird in der vorliegenden Arbeit vorgestellt.

Die genaue Kenntnis experimenteller Parameter erweist sich auch als entscheidender Faktor bei dem Bestreben, das Auflösungsvermögen von Elektronenmikroskopen in die Nähe von $1 \text{ \AA}$ voranzutreiben. Gegenwärtig arbeiten Gruppen in Deutschland (Universität Tübingen sowie Fa. Tietz, Gauting), den Niederlanden (Universität Delft sowie Fa. Philips, Eindhoven) und Belgien (Universität Antwerpen) im Rahmen eines von der Europäischen Gemeinschaft finanzierten Projektes, dem BRITE-EURAM project No. 3322, an diesem Vorhaben. Neben der instrumentellen Verbesserung der Elektronenmikroskope durch die Verwendung hochkohärenter Feldemissions-Elektronenquellen [5] und einem massiven Einsatz von digitalen Bilderfassungs- und Bildverarbeitungstechniken [6] hat sich die Strategie durchgesetzt, die Austrittswellenfunktion zu rekonstruieren.

Die Rekonstruktion der Austrittswellenfunktion bedeutet eine „Invertierung“ des elektronenoptischen Abbildungsprozesses, der zusätzlich zu den im Objekt stattfindenden dynamischen Beugungseffekten für die schwierige Interpretierbarkeit hochaufgelöster Aufnahmen verantwortlich ist. Ein Vorteil dieses Verfahrens besteht darin, daß die Austrittswellenfunktion frei von gerätespezifischen Einflüssen ist und abgesehen von der Wellenlänge des einfallenden Elektronenstrahls ausschließlich Objektinformation trägt. Im Gegensatz zu herkömmlichen elektronenmikroskopischen Aufnahmen, die aus gerätetechnischen Gründen Informationslücken bezüglich der Objektstruktur aufweisen, ist der Informationsgehalt der rekonstruierten Wellenfunktion im Idealfall lückenlos bis hinunter zum sogenannten Informationslimit. Das Informationslimit entspricht dem kleinsten Objektdetail, das in einer hochaufgelösten Abbildung noch als irgendwie geartetes Signal detektierbar ist.

Zwei Methoden werden gegenwärtig im Rahmen des BRITE-EURAM Projektes zur Rekonstruktion der Austrittswellenfunktion herangezogen. Die Methode der Elektronenholographie [7, 8], die eine instrumentelle Modifikation eines herkömmlichen Elektronenmikroskops erfordert, ist eher eine direkte Meßmethode als eine Rekonstruktionsmethode, da die in einem Hologramm gespeicherte Amplituden- und Phaseninformation zumindest theoretisch sehr einfach zu entschlüsseln ist. Bei der sogenannten Fokusvariationsmethode werden ca. 40–60 unter verschiedenen Fo-

kussierungsbedingungen aufgenommene herkömmliche Aufnahmen benutzt, um die Wellenfunktion zu rekonstruieren [9, 10, 11]. Bei dieser Methode wird die gesuchte Austrittswellenfunktion in numerisch aufwendiger Weise von den Einzelbildern aus „zurückgerechnet“.

In der vorliegenden Arbeit wird ein alternatives Verfahren zur Rekonstruktion der Austrittswellenfunktion aus herkömmlichen elektronenmikroskopischen Aufnahmen vorgestellt. Die „Invertierung“ des elektronenoptischen Abbildungsprozesses wird hierbei durch den Einsatz stochastischer Algorithmen erreicht. Dieses Verfahren eignet sich für die Rekonstruktion der zu periodisch aufgebauten Objekten gehörenden Wellenfunktionen. Es wird in der vorliegenden Arbeit anhand hochaufgelöster Abbildungen der intermetallischen Legierung Ni_4Mo und des Hochtemperatursupraleiters $\text{YBa}_2\text{Cu}_3\text{O}_7$ demonstriert.

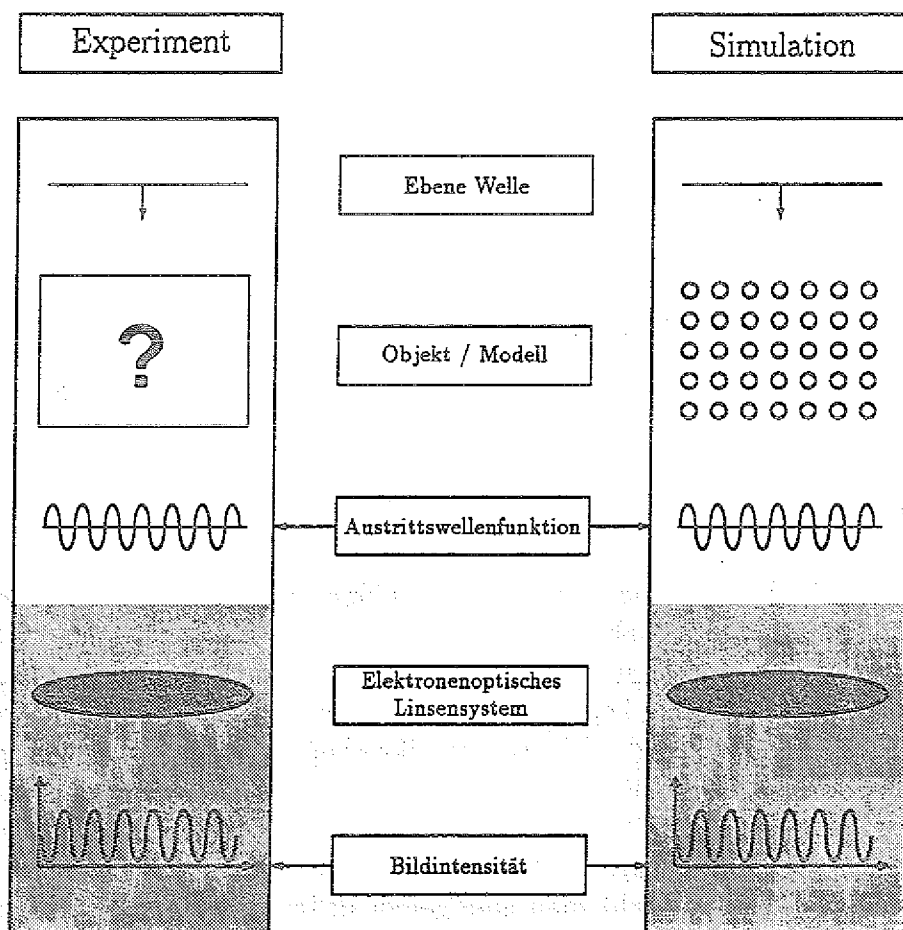


Abbildung 1.1: Schematische Darstellung der kontrasterzeugenden Prozesse in der hochauflösenden Elektronenmikroskopie. Die begleitende Bildsimulation, in der die experimentellen Vorgänge numerisch nachgebildet werden, ist parallel dargestellt. Hell unterlegt ist der dynamische Elektronenbeugungsprozeß, bei dem eine einlaufende ebene Elektronenwelle beim Durchqueren eines Objekts in Amplitude und Phase moduliert wird. Resultat dieses Beugungsprozesses ist die sogenannte Austrittswellenfunktion an der Objektunterseite. In einer zweiten Stufe (grau unterlegt) wird die Austrittswellenfunktion mittels des elektronenoptischen Linsensystems zur Abbildung gebracht. Vergleiche zwischen dem experimentellen Objekt und dem in die Simulation eingehenden Strukturmodell sind in der Bildebene anhand der Bildintensität oder – nach Einsatz eines Rekonstruktionsverfahrens – an der Objektunterseite anhand der Austrittswellenfunktion möglich (horizontale Pfeile).

Kapitel 2

Modellstrukturen

2.1 Die Struktur von Ni₄Mo

Die Legierung Ni₄Mo geht bei langsamen Abkühlen unter die Ordnungstemperatur von 860°C von einem nahgeordneten Zustand in einen ferngeordneten über, dessen Struktur in der Kristallographie mit D1_a bezeichnet wird. Untersuchungen des Ordnungsverhaltens der Legierung Ni₄Mo wurden mittels Elektronenbeugung [12, 13, 14, 15], konventioneller Elektronenmikroskopie [16], hochauflösender Elektronenmikroskopie [17] und Röntgenbeugung [18] durchgeführt. Der ferngeordnete Zustand der Legierung dient für die elektronenmikroskopisch-methodischen Untersuchungen, die in dieser Arbeit durchgeführt werden, als Modellstruktur.

Unter Fernordnung versteht man eine genau definierte periodische Anordnung von Atomen verschiedener Sorte auf einem Grundgitter. Im Idealfall erstreckt sich diese Ordnung über den gesamten makroskopischen Kristall. Dieses Grundgitter, das auch als Matrix bezeichnet wird, ist im speziellen Fall das kubisch-flächenzentrierte Gitter des reinen Nickels. Legierungen wie Ni₄Mo werden auch als substitutionelle Legierungen bezeichnet, da bestimmte Matrixatome (hier Ni) durch eine andere Atomsorte (hier Mo) ersetzt werden, ohne daß die Matrix verändert wird. Die D1_a Struktur ist dadurch bestimmt, daß die Mo- Fremdatome untereinander ein tetragonal-raumzentriertes Gitter mit

$$\begin{aligned}a_{D1a} &= b_{D1a} = \frac{1}{2}\sqrt{10}a_{kfz} \\c_{D1a} &= c_{kfz}\end{aligned}$$

bilden, das bei Blickrichtung entlang der c-Achse um den Winkel φ mit

$$\varphi = \arctan \frac{1}{3} \approx 18.43^\circ$$

gegen das kubisch-flächenzentrierte Gitter gedreht ist (Siehe Abb. 2.1). Aufgrund des größeren Radius der Mo- Atome tritt allerdings eine leichte Aufweitung des Ni-Gitters von $a_{Ni} = 0.351$ nm auf $a_{kfz} = 0.361$ nm auf. Darüberhinaus führt die Einstellung der Fernordnung zu einer Deformation des kfz Gitters mit $c_{kfz}/a_{kfz} = 0.986$,

so daß das mit *kfz* indizierte Matrixgitter genau betrachtet tetragonal flächenzentriert ist. Die Gitterkonstanten der tetragonal raumzentrierten Überstrukturzelle von Ni₄Mo, die in den Abbildungen 2.1 und 2.2 dargestellt ist, sind $a = b = 0.5720$ nm, $c = 0.3564$ nm [19].

Ein Charakteristikum der D1_a Struktur, die der Raumgruppe Nr. 87 mit dem Hermann-Mauguin Symbol I4/m angehört [20], ist die Fünferperiodizität der Ebenenabfolge entlang der $\langle 110 \rangle$ Richtungen, die in Abbildung 2.3 dargestellt ist. Auf eine Ebene, die nur aus Molybdänatomen besteht, folgen jeweils im Abstand $d_{550} = 0.08$ nm vier weitere Niclebenen.

Bei der Darstellung von Atomkoordinaten, kristallographischer Richtungen und Millerscher Indizes ist es möglich, sowohl die Einheitszelle der Matrix als auch die der Überstrukturzelle als Basis zu wählen. In dieser Arbeit wird die Überstrukturzelle als Bezugssystem zugrundegelegt, eine Bezugnahme auf die Matrixzelle wird jeweils explizit erwähnt. Die Einheitszelle von Ni₄Mo besteht aus zehn Atomen, deren Koordinaten in Tabelle 2.1 aufgelistet werden. Bei der gewählten Schreibweise erhält man Atompositionen der Einheitszelle mit $x/a = 1$, $y/b = 1$ oder $z/c = 1$ durch periodische Fortsetzung. Als Debeye-Waller-Faktor wurde in den Simulationen für Ni ein Wert von $B=0.0025$ nm² und für Mo ein Wert von $B=0.0035$ nm² benutzt [21].

Atom	x/a	y/b	z/c
Mo	0	0	0
Mo	0.5	0.5	0.5
Ni	0.4	0.2	0
Ni	0.8	0.4	0
Ni	0.6	0.8	0
Ni	0.2	0.6	0
Ni	0.1	0.3	0.5
Ni	0.3	0.9	0.5
Ni	0.7	0.1	0.5
Ni	0.9	0.7	0.5

Tabelle 2.1: Atompositionen der Ni₄Mo Einheitszelle.

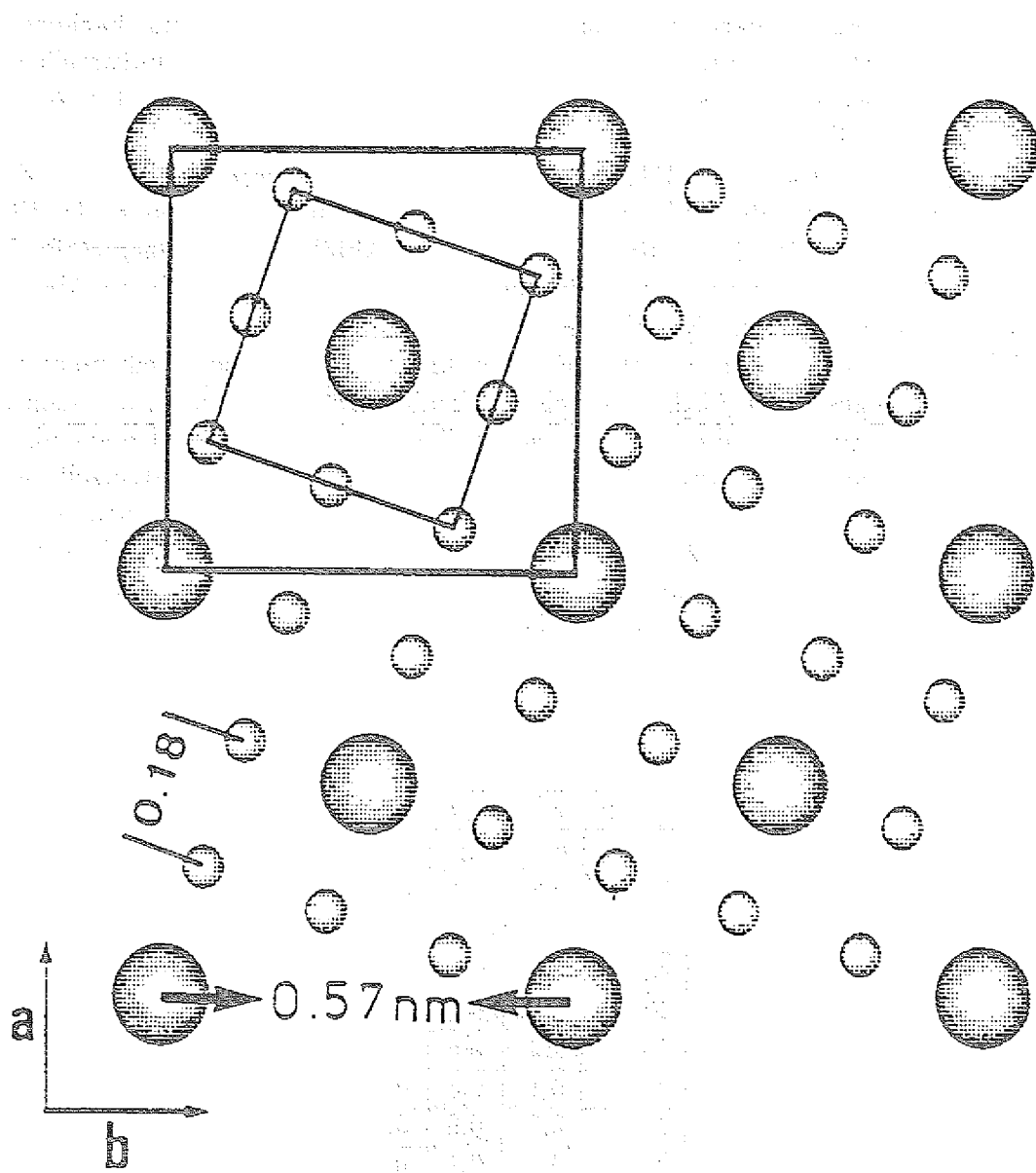


Abbildung 2.1: Projektion der D1_a Struktur entlang [001]. Die kfz Zelle ist mit dünnen Strichen gekennzeichnet, die D1_a Überstrukturzelle mit dicken. Große Kreise bezeichnen Mo, kleine bezeichnen Ni.

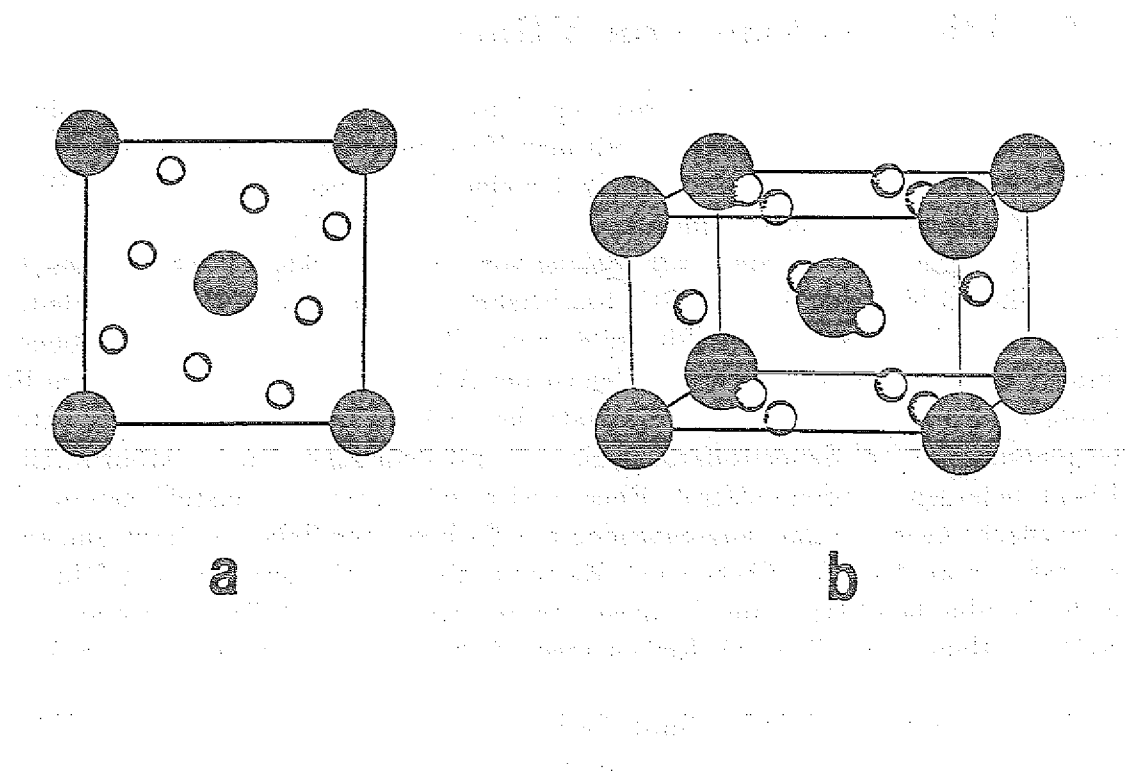


Abbildung 2.2: a) Projektion der D1_a Einheitszelle entlang $[001]$, b) Perspektivische Ansicht. Große Kreise bezeichnen Mo, kleine Kreise Ni.

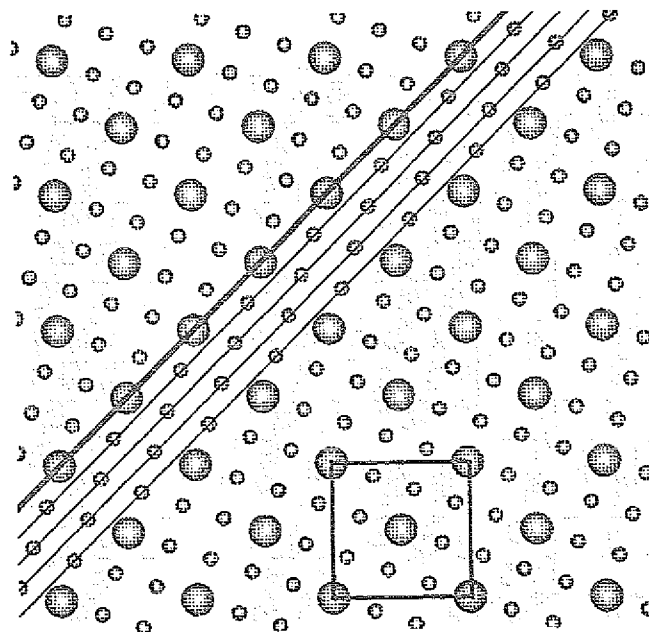


Abbildung 2.3: Darstellung der Fünferperiodizität in der Besetzungsabfolge der $\{550\}$ Ebenen. Der Abstand zwischen zwei $\{550\}$ Ebenen beträgt 0.08 nm . Eine Ni_4Mo Einheitszelle ist eingezeichnet.

2.2 Die Struktur von $\text{YBa}_2\text{Cu}_3\text{O}_7$

Die Einheitszelle des Hochtemperatursupraleiters $\text{YBa}_2\text{Cu}_3\text{O}_7$ ist orthorhombisch und gehört der Raumgruppe Nr. 47 mit dem Hermann-Mauguin Symbol Pmmm an [22]. Die Gitterkonstanten a, b, c nehmen bei einer Temperatur von 320 K die Werte $a = 0.38177 \text{ nm}$, $b = 0.38836 \text{ nm}$ und $c = 1.16827 \text{ nm}$ an [23].

Die Struktur läßt sich aus der Stapelung von drei Perowskit-ähnlichen Zellen, die durch die Stöchiometrieformel ABO_3 beschrieben werden, folgendermaßen herleiten: In allen drei Zellen sind die B-Gitterplätze mit Cu-Atomen besetzt. In der oberen und unteren der drei gestapelten Zellen ist der A-Gitterplatz von jeweils einem Ba-Atom besetzt, während der A-Gitterplatz der zentralen Zelle von einem Y-Atom eingenommen wird. Sauerstoffatome die sich mit dem zentralen Y-Atom in einer Ebene befinden, werden entfernt. Ebenso wird jedes zweite Sauerstoffatom an der Grenzfläche zwischen zwei angrenzenden mit Ba besetzten Zellen entfernt und zwar so, daß sich an der Grenzfläche CuO-Ketten ausbilden. Der letzte Schritt führt zur orthorhombischen Verzerrung der ansonsten tetragonalen Einheitszelle, wobei die b-Achse, entlang derer die CuO-Ketten angeordnet sind, etwas länger als die a-Achse ist.

Eine perspektivische Darstellung der Einheitszelle ist in Abbildung 2.4 zu sehen. Projektionen der Einheitszelle entlang den [100] und [010] Kristallrichtungen sind in Abbildung 2.5 dargestellt. Die Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ besteht aus 13 Atomen, deren Koordinaten mit den zugehörigen Besetzungszahlen und Debeye-Waller-Faktoren in Tabelle 2.2 aufgelistet sind [24]. (Atompositionen der Einheitszelle mit $x/a = 1$, $y/b = 1$ oder $z/c = 1$ ergeben sich durch periodische Fortsetzung.)

Atom	x/a	y/b	z/c	Besetzung	D.-W.-F. [nm^2]
Y	0.5	0.5	0.5	1.00	0.0049
Ba	0.5	0.5	0.1844	1.00	0.0064
Ba	0.5	0.5	0.8156	1.00	0.0064
Cu1	0	0	0	1.00	0.0051
Cu2	0	0	0.3552	1.00	0.0052
Cu2	0	0	0.6448	1.00	0.0052
O1	0	0	0.1588	1.00	0.0075
O1	0	0	0.8412	1.00	0.0075
O2	0.5	0	0.3779	1.00	0.0062
O2	0.5	0	0.6221	1.00	0.0062
O3	0	0.5	0.3780	1.00	0.0064
O3	0	0.5	0.6220	1.00	0.0064
O4	0	0.5	0	0.91	0.0050

Tabelle 2.2: Strukturdaten von $\text{YBa}_2\text{Cu}_3\text{O}_7$.

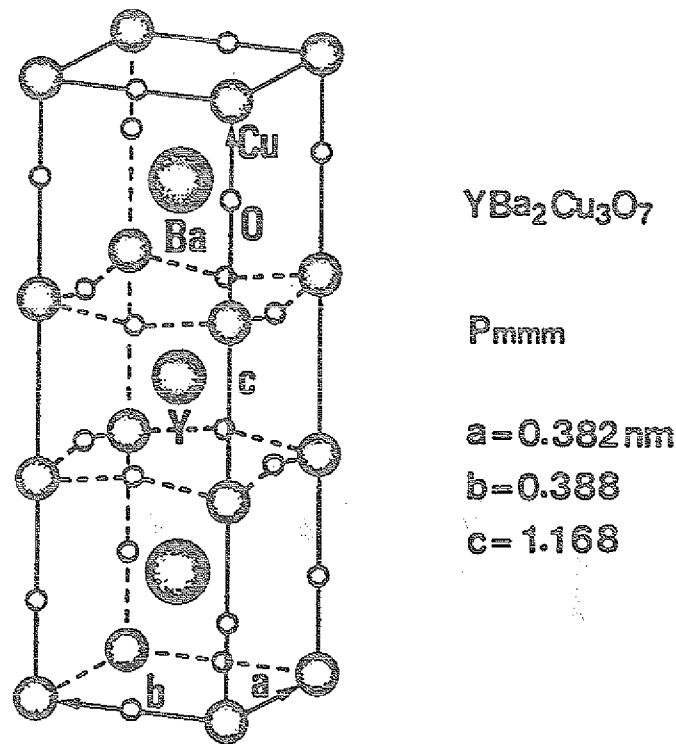


Abbildung 2.4: Perspektivische Darstellung einer Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$.

Je nach Herstellungsprozeß tritt in den entlang der b -Achse verlaufenden CuO -Ketten eine Sauerstoffdefizienz auf, die sich in einer von 1 abweichenden Besetzungszahl der $\text{O}4$ -Position ausdrückt. Entsprechend kann die Anzahl der Sauerstoffatome pro Einheitszelle zwischen 6 (alle $\text{O}4$ -Positionen unbesetzt) und 7 (alle $\text{O}4$ -Positionen besetzt) variieren. Die Sprungtemperatur T_c , bei der Supraleitung eintritt, liegt nur dann über der Siedetemperatur von flüssigem Stickstoff, falls die Besetzungszahl der $\text{O}4$ -Position nahe 1 liegt [25].

In den letzten Jahren ist eine große Anzahl von Publikationen erschienen, in denen elektronenmikroskopisch hochaufgelöste Abbildungen von $\text{YBa}_2\text{Cu}_3\text{O}_7$ eine wichtige Rolle spielen. Dabei standen aufgrund des hohen technologischen Potentials der Hochtemperatursupraleitung hauptsächlich materialwissenschaftliche Aspekte im Vordergrund. In diesem Zusammenhang wurden insbesondere dünne Filme aus $\text{YBa}_2\text{Cu}_3\text{O}_7$ untersucht, die epitaktisch auf Substraten wie MgO , SrTiO_3 oder LaAlO_3 deponiert wurden [26, 27, 28, 29, 30, 31, 32]. Untersuchungen, bei denen die methodischen Aspekte der hochauflösenden Abbildung von $\text{YBa}_2\text{Cu}_3\text{O}_7$ im Vordergrund stehen, sind eher selten [33, 34, 35, 36, 37, 38].

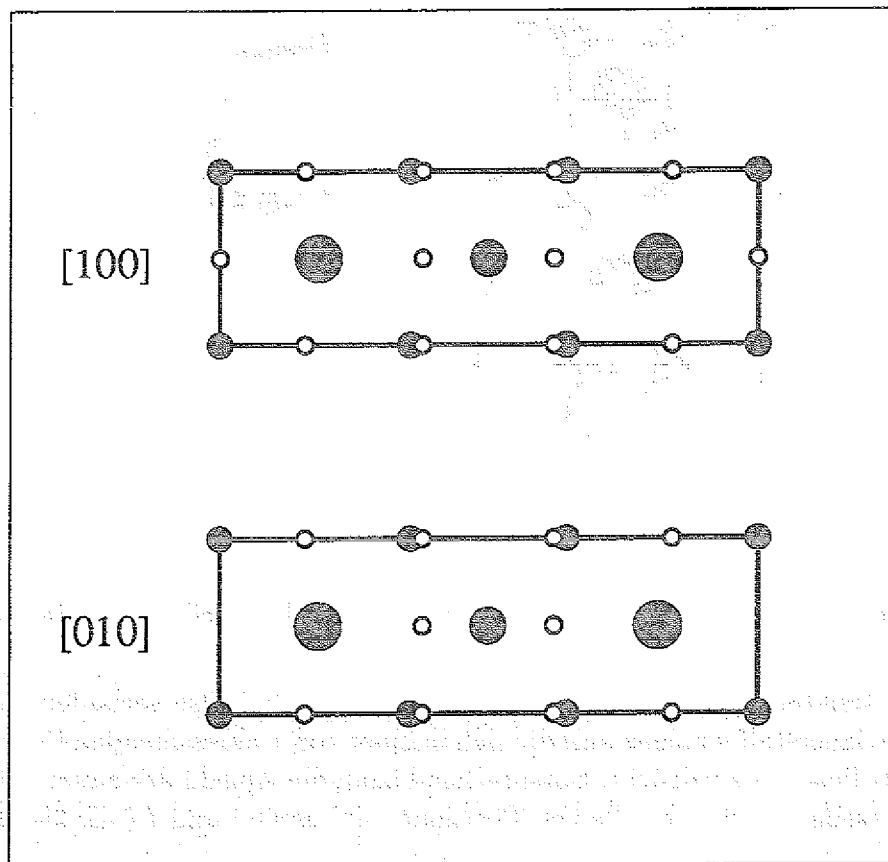


Abbildung 2.5: Oben: Projektion der Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ entlang der $[100]$ Kristallrichtung. Unten: Projektion entlang der $[010]$ Kristallrichtung. Volle, große Kreise bezeichnen Ba, mittlere Y und kleine Cu. Sauerstoffpositionen sind durch offene Kreise markiert.

Kapitel 3

Grundlagen hochauflösender Abbildung

3.1 Aufbau eines Transmissionselektronenmikroskops

Moderne Transmissionselektronenmikroskope bestehen aus bis zu neun magnetischen Linsen und einer Vielzahl von Ablenkspulen. Zur Steuerung dieser Elemente kommen komplexe elektronische Bauteile zum Einsatz. Der Aufbau eines Mikroskops läßt sich jedoch bei prinzipieller Betrachtungsweise durch einige wenige Begriffe erklären (Siehe Abb. 3.1).

- Elektronenkanone
Elektronen, die durch Glühemission aus einer Kathode, dem sogenannten *Filament*, austreten, durchlaufen eine Beschleunigungsstrecke. Die Elektronen erlangen dabei Energien von typischerweise einigen hundert keV.
- Kondensorsystem
Aus den aus einem Beschleuniger austretenden Elektronen wird vom Kondensorsystem, das aus bis zu drei Linsen und einer darauffolgenden Blende besteht, ein definierter Strahl erzeugt.
- Präparat
Das Präparat muß für die einfallenden Elektronen weitgehend transparent sein. Der einfallende Elektronenstrahl wird durch die Wechselwirkung der Elektronen mit den streuenden Potentialen der im Präparat befindlichen Atome abgelenkt.
- Objektivlinse
 - Abgelenkte Strahlen, die das Präparat unter gleichen Austrittswinkeln verlassen haben, werden von der Objektivlinse in ihrer hinteren Brennebene zum sogenannten *Beugungsbild* vereinigt.

- Nahe der dahinterliegenden zur Lichtoptik analogen Gaußschen Bildebene entsteht das erste *Zwischenbild* der Präparatstruktur.
- **Objektivblende**
Sie liegt nahe der hinteren Brennebene der Objektivlinse. Mit der Objektivblende kann die Anzahl der abgebeugten Strahlen, die zur Bildentstehung beitragen sollen, begrenzt werden. Die Anzahl inelastisch gestreuter Elektronen, die nicht zum Bildkontrast beitragen, wird dabei ebenso reduziert.
- **Selektorblende**
Die Selektorblende liegt in der Ebene des ersten Zwischenbildes. Mit ihr kann im Beugungsmodus (s.u.) der zum Beugungsbild beitragende Teil des Präparats räumlich begrenzt werden.
- **Zwischenlinsen**
Sie dienen zur Nachvergrößerung entweder des
 - Zwischenbildes (siehe Abb. 3.1 (a)) oder des
 - Beugungsbildes (siehe Abb. 3.1 (b))

Im ersten Fall spricht man vom Abbildungsmodus im letzteren vom Beugungsmodus des Geräts.

- **Projektorlinse**
Sie wirft ein nocheinmal nachvergrößertes Bild auf einen Fluoreszenzschirm. Dieses Bild kann auch auf Photoplatte aufgenommen werden oder – je nach Ausstattung – mittels einer Video- oder einer CCD- (*charge coupled device*) Kamera betrachtet und digital gespeichert werden.

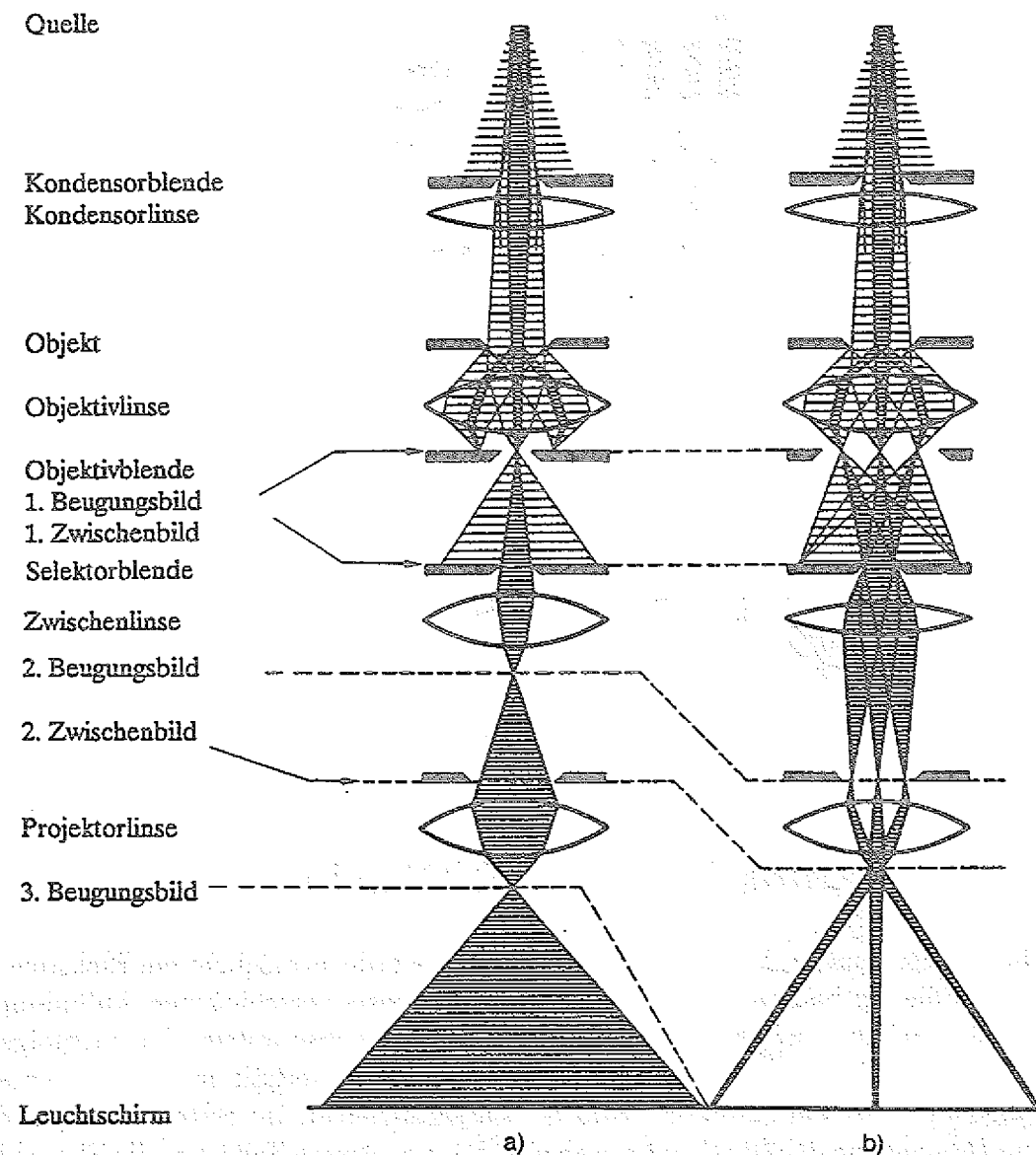


Abbildung 3.1: (a) Strahlengang eines Elektronenmikroskops im Abbildungsmodus. (b) Strahlengang im Beugungsmodus. Der Übersichtlichkeit halber ist jeweils nur eine Kondensor- und eine Zwischenlinse eingezeichnet.

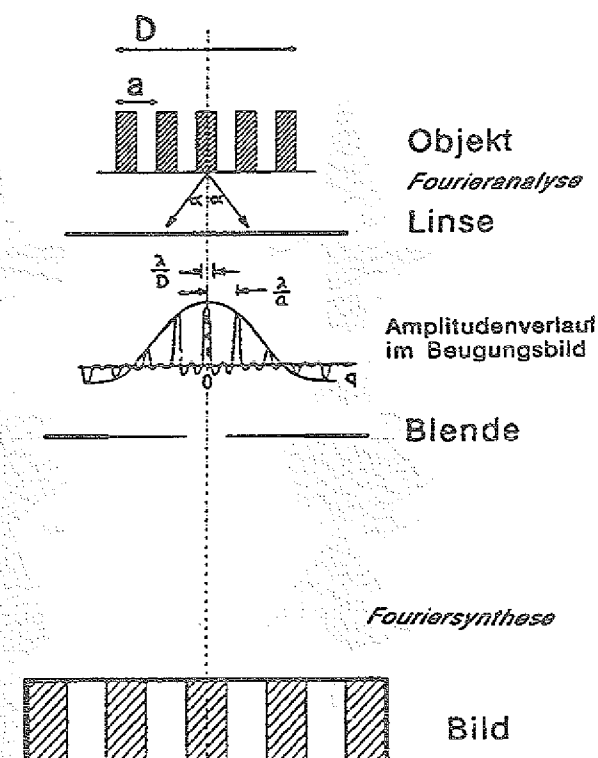


Abbildung 3.2: Darstellung des Prinzips hochauflösender Abbildung an einem eindimensionalen Objekt der räumlichen Periode a .

3.2 Abbildung kristalliner Objekte

Wie im vorangehenden Abschnitt bereits erwähnt wurde, ermöglicht ein Elektronenmikroskop die Beobachtung der Objektstruktur in zwei verschiedenen Abbildungsmoden. Je nachdem, welche Ebene hinter der Objektivlinse mittels der nachfolgenden Linsen auf dem Fluoreszenzschirm abgebildet wird, erhält man entweder ein Beugungsbild oder ein „direktes“ Bild der Objektstruktur. Im ersten Fall wird die hintere Brennebene der Objektivlinse abgebildet, im zweiten Fall ist es die Ebene des ersten Zwischenbildes. Genau genommen wird nicht das Objekt selbst, sondern die aus dem Objekt austretende Elektronenwelle zur Abbildung gebracht. Hierbei entspricht die Wahl des Beugungsmodus einer Fourieranalyse der resultierenden Elektronenwelle, die Wahl des Abbildungsmodus entspricht einer nochmaligen Fouriertransformation, die wieder zur Synthese der Welle führt.

3.2.1 Analyse der Objektstruktur

Grundlegende Prinzipien der Elektronenbeugung seien anhand eines eindimensionalen Objekts der räumlichen Periode a erklärt (Siehe Abbildung 3.2).

Eine ebene Elektronenwelle mit Wellenvektor $\vec{k}_0$, die den einfallenden Strahl beschreibt, wird an der Objektstruktur, die eine räumliche Periodizität der Länge a

besitzt, gebeugt. Die einfallende Welle wird beschrieben durch:

$$\Psi_0 = A_0 \exp(2\pi i \vec{k}_0 \vec{r}). \quad (3.1)$$

Analog zur Beugung eines Lichtstrahls an einem optischen Gitter verlassen neben dem durchgehenden Strahl mit dem Wellenvektor $\vec{k}_0$ nun auch unter Winkeln α_n abgebeugte Strahlen mit den zugehörigen Wellenvektoren $\vec{k}_n$ das Objekt:

$$\Psi_n = A_n \exp(2\pi i \vec{k}_n \vec{r}). \quad (3.2)$$

Die Bedingung für die Existenz eines abgebeugten Strahls ist durch die sogenannte *Lauebedingung* gegeben:

$$\vec{k}_n - \vec{k}_0 = \vec{g}_n. \quad (3.3)$$

Im Falle des eindimensionalen Beispiels aus Abb. 3.2 gilt für $|\vec{g}_n|$:

$$|\vec{g}_n| = n/a \quad \text{mit } n = \pm 1, \pm 2, \dots \quad (3.4)$$

Die *Lauebedingung* (3.3), die bei hochauflösenden Abbildungen – wie später dargelegt wird – streng genommen nie ganz erfüllt ist, ordnet einem abgebeugten Strahl

$$\Psi_n = A_n \exp(2\pi i \vec{g}_n \vec{r}) \quad (3.5)$$

die räumliche Periodizität $1/|\vec{g}_n| = a/n$ der Objektstruktur zu. Der dem einfallenden Strahl und allen abgebeugten Strahlen gemeinsame Phasenfaktor $\exp(2\pi i \vec{k}_0 \vec{r})$ wird wie in Gleichung (3.5) üblicherweise weggelassen.

Geht man von einer beliebigen periodischen Objektstruktur aus, läßt sich das Wellenfeld nach Verlassen des Objekts als Summe über die einzelnen Teilwellen Ψ_n beschreiben:

$$\Psi(\vec{r}) = \sum_n \Psi_n = \sum_n A_n \exp(2\pi i \vec{g}_n \vec{r}). \quad (3.6)$$

Die Funktion in Gleichung (3.6) wird *Austrittswellenfunktion* (engl.: *exit surface wavefunction*) genannt. Sie enthält die gesamte Information über die Wechselwirkung des einfallenden Elektronenstrahls mit dem Objekt. Vom mathematischen Standpunkt ist die Austrittswellenfunktion eine Fouriersumme über die einzelnen Teilwellen Ψ_n . In Abbildung 3.3 ist die sukzessiv genauer werdende Darstellung der Objektfunktion aus Abbildung 3.2 durch die Einbeziehung höherer Ordnungen ($n = 1, n = 3, n = 5$) in die Reihenentwicklung (3.6) dargestellt.

Die Objektivlinse vereinigt nun alle unter demselben Winkel austretenden Strahlen in ihrer hinteren Brennebene in einem Punkt. Dieser Vorgang entspricht einer Zerlegung der Austrittswellenfunktion $\Psi(\vec{r})$ in die Teilwellen Ψ_n , also einer *Fourieranalyse*.

- Die Intensität I_n der Beugungspunkte, die auf dem Fluoreszenzschirm beobachtet werden können, entspricht dem Betragsquadrat der komplexen Fourierkoeffizienten A_n :

$$I_n = |A_n|^2. \quad (3.7)$$

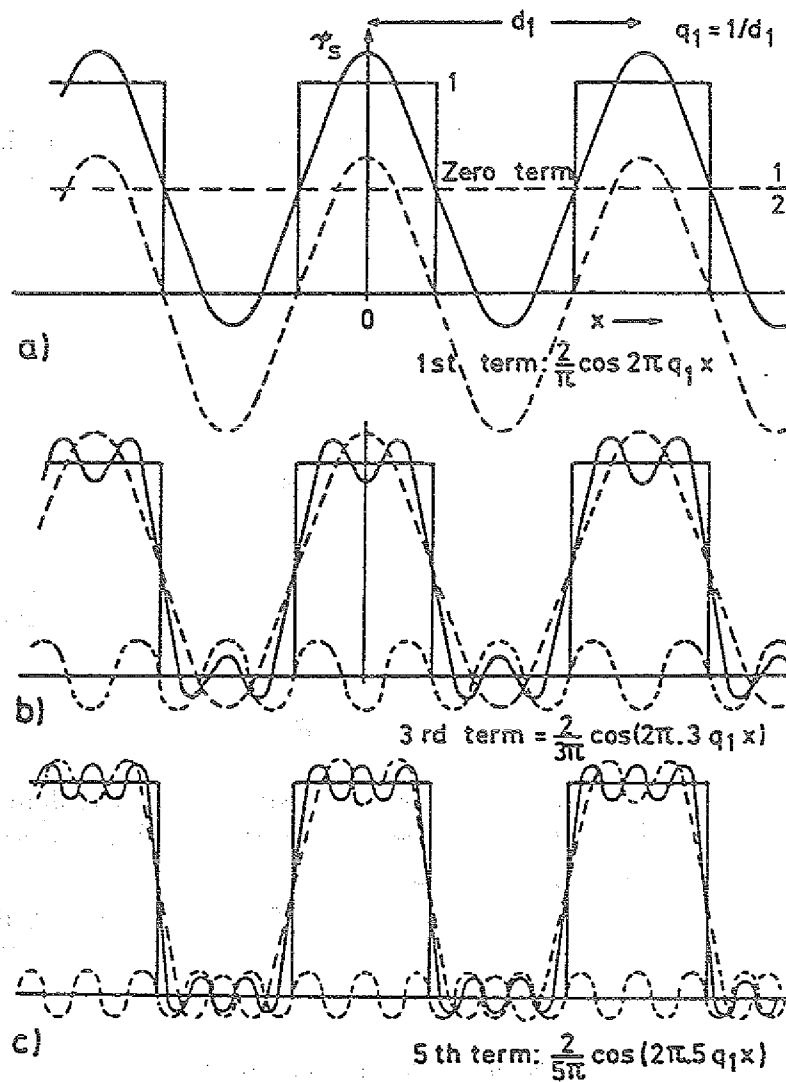


Abbildung 3.3: Darstellung der sukzessive genauer werdenden Beschreibung einer Rechteckfunktion durch Einbeziehung von Fourierkomponenten höherer Ordnung.

Wegen der Betragsquadratbildung in Gleichung (3.7) läßt sich die Wellenfunktion aus dem Beugungsbild nicht mehr rekonstruieren, da die Phaseninformation der komplexen Koeffizienten A_n verlorengeht.

- Die Lage der Beugungspunkte im Beugungsbild ist durch die Winkel α_n zwischen dem einfallenden Strahl und den abgebeugten Strahlen festgelegt

$$\alpha_n \approx \sin \alpha_n = \lambda |\vec{g}_n|, \quad (3.8)$$

wobei λ die Wellenlänge des einfallenden Strahls ist. Der Abstand eines Beugungspunktes n -ter Ordnung vom Ursprung des Fourierraums ist gerade $|\vec{g}_n|$. Je kleiner also die räumliche Periode a , die das Objekt charakterisiert, ist, desto größer sind die zugeordneten *Raumfrequenzen* $|\vec{g}_n| = n/a$ und die entsprechenden Beugungswinkel α_n .

Geht man von dem bisherigen eindimensionalen Beispiel ab, stellt sich die Frage, welche Fourierkomponenten bei einem dreidimensionalen kristallinen Objekt unter gegebenen Einstrahlungsbedingungen eine Rolle spielen. Nach den obigen Ausführungen kann man die Frage auch anders formulieren: Welche Beugungspunkte sieht man bei einer bestimmten Probenorientierung auf dem Fluoreszenzschirm?

Das reziproke Gitter

Die in Lehrbüchern (z.B. [39]) im Rahmen der *kinematischen Beugungstheorie* ausführlich beschriebene Konstruktion der *Ewaldkugel* im reziproken Gitter sei hier kurz wiederholt:

Das Rezept zur Konstruktion der Ewaldkugel (Abb. 3.4), mit der die Vektoren $\vec{g}$ im Fourierraum, die die Lauebedingung Gl. (3.3) erfüllen, ermittelt werden können, ist folgendes:

- Man konstruiere das reziproke Gitter der Einheitszelle (s.u.).
- Man wähle einen beliebigen reziproken Gitterpunkt als Ursprung O und ziehe eine Linie MO parallel zu $\vec{k}_0$, der Richtung des einfallenden Strahls, mit $|\vec{k}_0| = 1/\lambda$.
- Man zeichne einen Kreis mit Mittelpunkt M und Radius $|\vec{k}_0|$.
- Falls der Kreis einen reziproken Gitterpunkt G, dessen Koordinaten durch $\vec{g}$ gegeben sind, schneidet, ist die Lauebedingung erfüllt. Zusätzliche Voraussetzung für die Existenz einer abgebeugten Welle Ψ_n ist, daß die sogenannte *Strukturamplitude* $F(\vec{g}) \neq 0$ ist.

Abb. 3.4 verdeutlicht diese Konstruktion anhand eines zweidimensionalen Rechteckgitters mit den Gitterkonstanten a und b , wobei $a > b$ ist. In Abb. 3.4 (a) ist die Richtung der einfallenden Welle so gewählt, daß die Lauebedingung für den Punkt G exakt erfüllt ist. Die in Abb. 3.4 (b) dargestellte Situation entspricht einer hochsymmetrischen Einstrahlrichtung parallel zu b . Eine solche hochsymmetrische

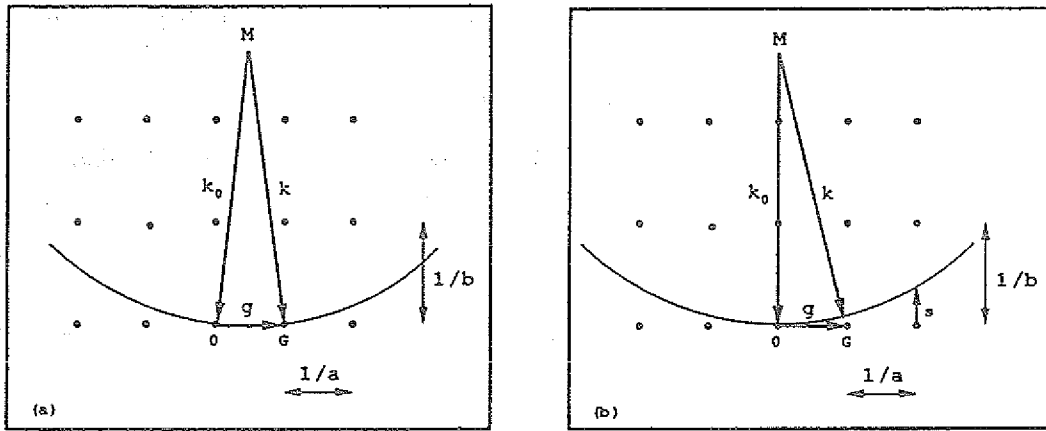


Abbildung 3.4: Konstruktion der Ewaldkugel an einem Rechteckgitter mit den Gitterkonstanten a und b im reziproken Raum. In (a) ist die Lauebedingung exakt erfüllt. Der Fall (b) entspricht einer hochsymmetrischen Einstrahlrichtung entlang einer der Rechteckachsen (hier b). Die Krümmung der Ewaldkugel ist stark übertrieben dargestellt.

Einstrahlrichtung wird üblicherweise in der hochauflösenden Elektronenmikroskopie benutzt. In diesem Falle ist die Lauebedingung für keinen Punkt des reziproken Gitters streng erfüllt. Die Abweichung von der Lauebedingung wird für jeden Punkt des reziproken Gitters bei gegebener Einstrahlungsrichtung durch den sogenannten *Anregungsfehler* $\vec{s}$ angegeben. In der *kinematischen Beugungstheorie* wird gezeigt, wie durch den Einfluß der sogenannten *Gitteramplitude* auch Strahlen angeregt werden, die die Lauebedingung *nicht* exakt erfüllen.

Die zur Konstruktion der Ewaldkugel benötigten reziproken Gitterpunkte lassen sich durch Linearkombination der reziproken Gittervektoren $\vec{a}_i^*$ darstellen:

$$\vec{g} = h\vec{a}_1^* + k\vec{a}_2^* + l\vec{a}_3^*. \quad (3.9)$$

Die reziproken Gittervektoren erhält man durch folgende Beziehung:

$$\vec{a}_1^* = \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)}, \quad (3.10)$$

wobei die $\vec{a}_i$ die Gittervektoren der Einheitszelle sind. Für $\vec{a}_2^*$ und $\vec{a}_3^*$ erhält man die entsprechende Formel durch Permutation der Indizes.

Für die Hochauflösung praktikable Einfallrichtungen des Elektronenstrahls, die einer hochsymmetrischen Orientierung des Kristalls entsprechen, lassen sich darstellen durch (Siehe Abb. 3.5 (a)):

$$\vec{r} = m\vec{a}_1 + n\vec{a}_2 + o\vec{a}_3. \quad (3.11)$$

Alle reziproken Gittervektoren $\vec{g} = h\vec{a}_1^* + k\vec{a}_2^* + l\vec{a}_3^*$, die die Bedingung

$$mh + nk + ol = 0 \quad (3.12)$$

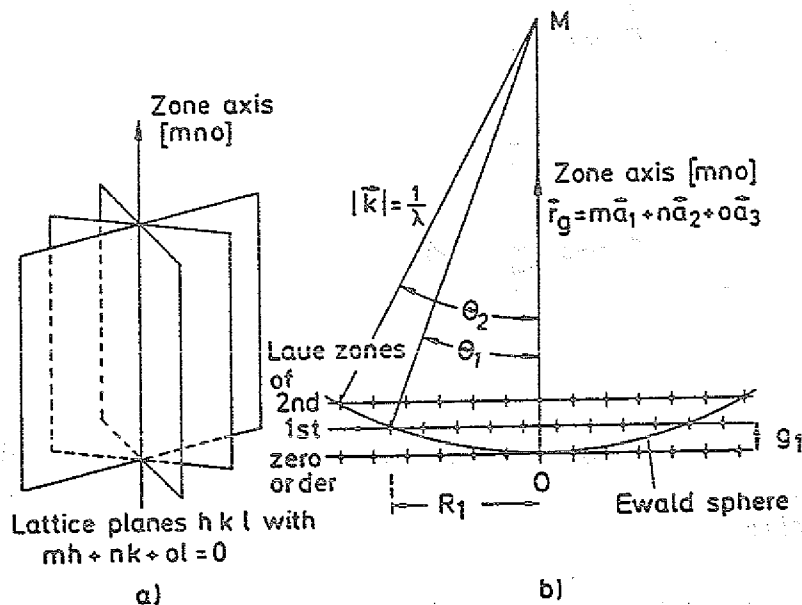


Abbildung 3.5: (a) Gitterebenen mit gemeinsamer Zonenachse [mno]. (b) Lage der nullten Lauezone im reziproken Gitter. Lauezonen höherer Ordnung sind ebenfalls eingezeichnet. Die Krümmung der Ewaldkugel ist stark übertrieben dargestellt.

erfüllen, gehören zur *Lauezone nullter Ordnung* des reziproken Gitters. Der Vektor $\vec{r}$ aus Gleichung (3.11) wird auch Zonenachse genannt.

Die reziproken Gitterpunkte der nullten Lauezone sind diejenigen, die bei Einstrahlung entlang der Zonenachse am ehesten die Lauebedingung erfüllen, also den kleinsten Anregungsfehler besitzen. Dies liegt daran, daß bei Elektronenenergien von einigen hundert keV der Radius der Ewaldkugel in den meisten Materialien typischerweise um zwei Größenordnungen größer ist als der Abstand zur nächst höheren Lauezone. Man kann die Ewaldkugel also im Falle hoher Elektronenenergien durch eine Ebene annähern.

Dies bedeutet, daß die gesuchten Gittervektoren $\vec{g}$ aus der Fourierentwicklung der Austrittswellenfunktion einfach durch die reziproken Gitterpunkte bestimmt sind, die auf einer durch den Koordinatenursprung gehenden senkrecht zur Einfallsrichtung liegenden Ebene des reziproken Gitters liegen.

In Abb. 3.6 (a) ist das reziproke Gitter von Ni_4Mo dargestellt. Die großen Kreise geben reziproke Gitterpunkte an, die bei statistisch zufälliger Verteilung von Ni- und Mo-Atomen auf dem Matrixgitter existieren würden. Dies ist äquivalent zu dem Fall, daß man im Realraum einen kfz-Kristall hat, bei dem sich auf jedem Gitterpunkt ein im Verhältnis 4:1 aus Ni und Mo „gemischtes“ Atom befindet. Man erhält also im Falle eines *ungeordneten* Kristalls dieselben reziproken Gitterpunkte, die auch das kfz-Gitter aufweist. Diese Gitterpunkte sind deshalb in Abb. 3.6 auch mit reziproken Koordinaten versehen, die sich auf das kfz-Gitter beziehen.

Zusätzlich treten – im Falle des *ferngeordneten* Kristalls – als Ausdruck der Periodizität in der Besetzung mit den beiden Atomsorten noch sogenannte Über-

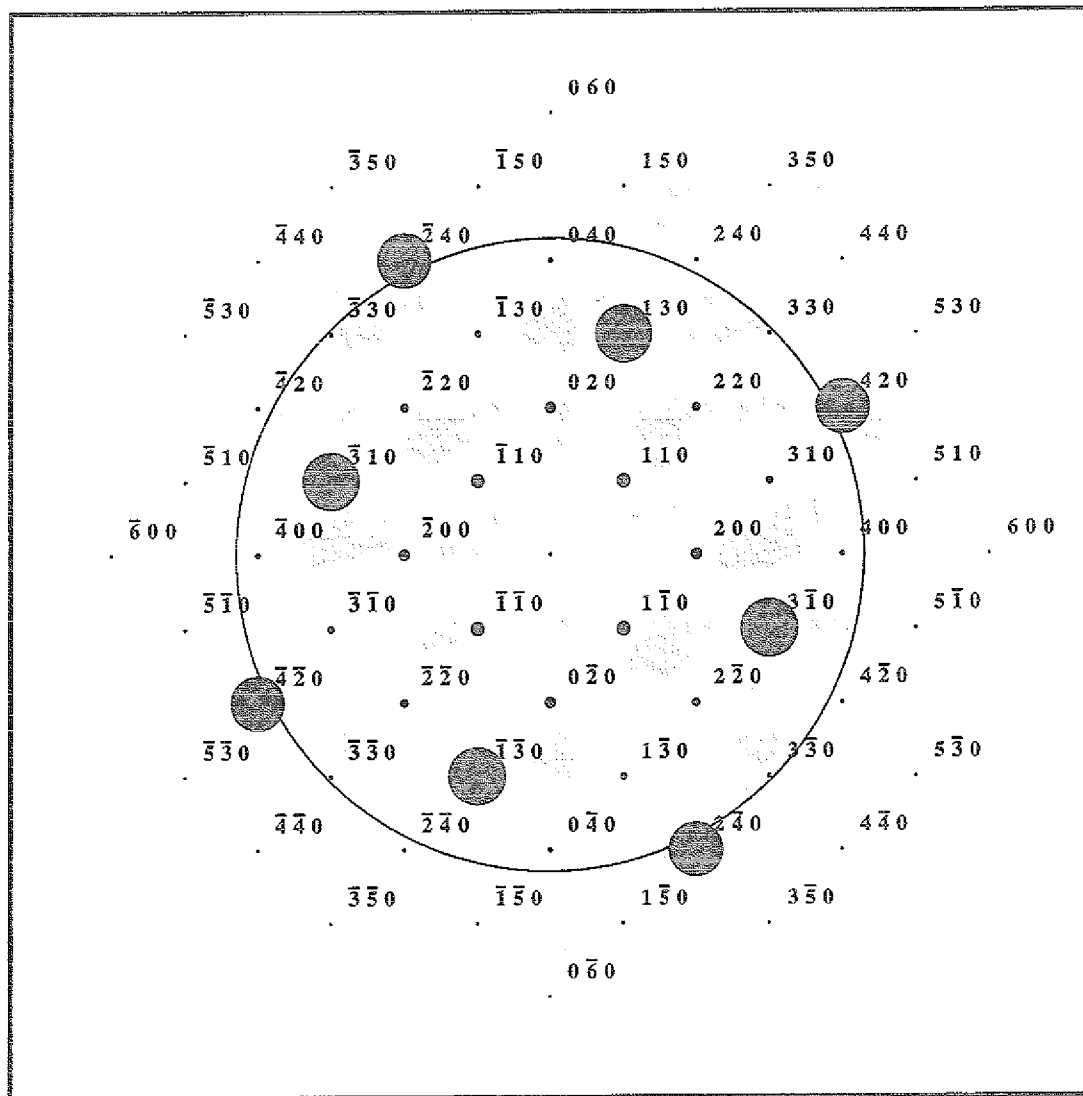


Abbildung 3.7: Nach der kinematischen Theorie berechnetes Beugungsbild von Ni_4Mo bei Einstrahlrichtung entlang $[001]$. Abgesehen vom Nullstrahl ist der Durchmesser der Punkte proportional zur berechneten Intensität. Die Indizes (hkl) beziehen sich auf die D1_a Einheitszelle. Der eingezeichnete Kreis entspricht einer Objektivblende mit einem reziproken Radius von 7.5 nm^{-1} . In diesem Falle befinden sich 29 Strahlen innerhalb der Blende (Siehe Kapitel 8).

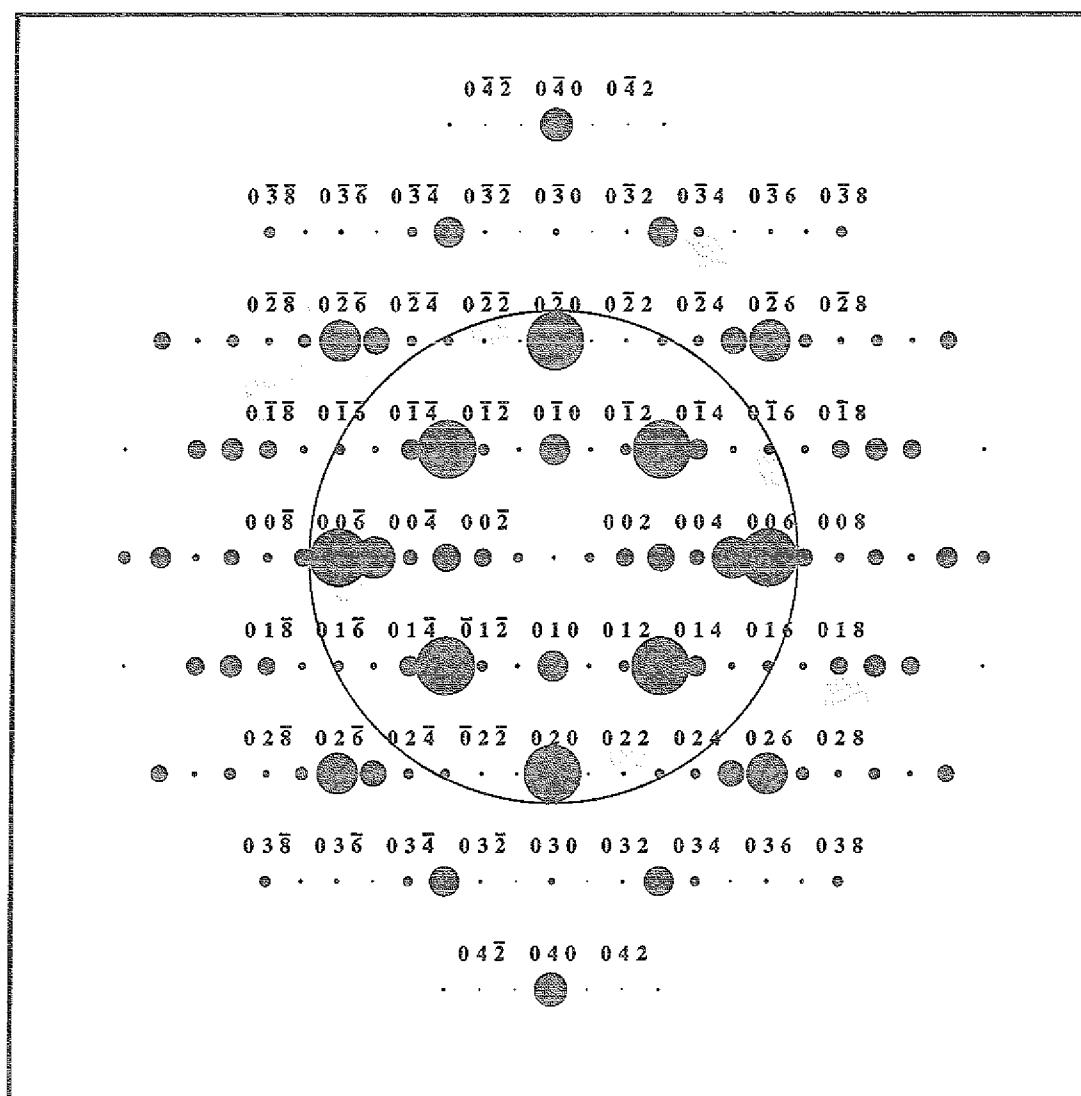


Abbildung 3.8: Nach der kinematischen Theorie berechnetes Beugungsbild von $\text{YBa}_2\text{Cu}_3\text{O}_7$ bei Einstrahlrichtung entlang $[100]$. Abgesehen vom Nullstrahl ist der Durchmesser der Punkte proportional zur berechneten Intensität. Aus Platzgründen sind nur Reflexe $(0kl)$ mit geradem l indiziert. Der eingezeichnete Kreis entspricht einer Objektivblende mit einem reziproken Radius von 5.5 nm^{-1} . In diesem Falle befinden sich 53 Strahlen innerhalb der Blende (Siehe Kapitel 8).

wird. Sie sind für jedes chemische Element unterschiedlich und hängen auch vom Bindungszustand ab. Die aus relativistischer *Hartree-Fock*-Berechnung gewonnenen Werte für die atomaren Streuamplituden freier Atome können Tabellen [40, 41] entnommen werden.

In Abb. 3.7 ist das kinematisch berechnete Beugungsbild von Ni_4Mo bei Einstrahlrichtung entlang [001] dargestellt. Der Durchmesser der Beugungspunkte ist hierbei ein Maß für die berechnete Intensität. Es ist klar zu erkennen, daß die Strukturamplituden der Überstrukturreflexe dem Betrage nach deutlich kleiner sind als die derjenigen Reflexe, die von der kfz-Matrix herrühren. In Abb. 3.8 ist das kinematisch berechnete Beugungsbild von $\text{YBa}_2\text{Cu}_3\text{O}_7$ bei Einstrahlrichtung entlang der [100] Zonenachse zu sehen. Die dicht gedrängte Abfolge von Beugungspunkten (hkl) und ($hkl \pm 1$) ist auf die relativ lange c-Achse von $\text{YBa}_2\text{Cu}_3\text{O}_7$ zurückzuführen. Die Wellenvektoren der Reflexe {020}, {006} und {013} verlaufen etwa parallel zu den Kanten und den Diagonalen der Perowskit-Grundzellen, aus denen die $\text{YBa}_2\text{Cu}_3\text{O}_7$ Struktur abgeleitet werden kann. In weiterem Sinne können diese stark angeregten Reflexe deshalb auch als Matrixreflexe betrachtet werden.

3.2.2 Synthese der Objektstruktur

Gemäß dem Huygens-Fresnelschen Prinzip wirken die zu den einzelnen Raumwellenlängen gehörenden Beugungspunkte als Punktquellen für Kugelwellen. Diese Kugelwellen erzeugen in der Bildebene der Objektlinse ein Interferenzbild, das im Falle einer *idealen* Linse dem Betragsquadrat der Austrittswellenfunktion $\Psi(\vec{r})$ entspricht: (Der Einfachheit halber wird mit der Vergrößerung $M = 1$ gerechnet.)

$$I_{\text{Bild}}^{\text{ideal}}(\vec{r}) = |\Psi(\vec{r})|^2. \quad (3.15)$$

In etwas ungenauer Formulierung kann man hier von einer zweiten Fouriertransformation sprechen, die einer Rücktransformation des Beugungsbildes zur Realstruktur des Objekts entspricht (Abb. 3.2). Diese Fouriersynthese ist jedoch mit „Fehlern“ behaftet, die zum einen in der Abbildungsmethode zu suchen sind, zum anderen durch die Grenzen des technisch Machbaren verursacht sind. Diese Unzulänglichkeiten, die sowohl die Bildinterpretation erschweren als auch die Auflösung begrenzen, werden im folgenden aufgezählt:

Nichtlinearität der Abbildung

Durch die Betragsquadratbildung in Gleichung (3.15) treten in der Abbildung sogenannte *Artefakte* auf. Mit Artefakt bezeichnet man allgemein Bildstrukturen, die nicht mehr unmittelbar im Sinne der intuitiven Bildinterpretation mit der Objektstruktur in Einklang zu bringen sind. Dies sei an einem ganz simplen Beispiel verdeutlicht:

Die Austrittswellenfunktion enthalte die beiden Fourierkomponenten

$$\Psi_{\pm\vec{g}} = A_g \exp(\pm 2\pi i \vec{g} \vec{r}). \quad (3.16)$$

Die konjugiert komplexe Austrittswellenfunktion enthält dann die entsprechenden Terme

$$\Psi_{\pm\vec{g}}^* = A_g^* \exp(\mp 2\pi i \vec{g}\vec{r}). \quad (3.17)$$

Im Produkt $\Psi\Psi^*$ aus Gleichung (3.15), die ein ideales Elektronenmikroskop beschreibt, tritt als Beitrag zur Bildintensität folgender Term auf:

$$I(\vec{r}) = \dots + 2|A_g|^2 [1 + \cos(4\pi \vec{g}\vec{r})] + \dots \quad (3.18)$$

Dieser Term beinhaltet neben einem ortsunabhängigen Anteil, der nicht zum Bildkontrast beiträgt, die Raumwellenlänge der Frequenz $2g$, obwohl in der zugrundeliegenden Wellenfunktion nur Raumwellenlängen der Frequenz g auftreten. Das Bild weist in diesem Falle also feinere Details als die Wellenfunktion auf.

Selektive Phasenmodulation

Die einzelnen Teilwellen $\Psi_{\vec{g}}$ erfahren vor der Fouriersynthese aufgrund prinzipieller Unzulänglichkeiten magnetischer Linsen Phasenverschiebungen, die in starkem Maße vom Beugungsvektor $\vec{g}$ abhängen. Diese „Durchmischung“ der in den komplexen Fourierkoeffizienten A_g (Gl. 3.6) steckenden Phasenbeziehungen verändert nachhaltig und auf intuitiv nicht mehr nachvollziehbare Art und Weise die Wellenfunktion und somit auch das Bild.

Partielle Kohärenz

In den bisherigen Ausführungen wurde der auf das Objekt einfallende Strahl idealisiert als *monochromatische ebene* Welle beschrieben. Beide Idealisierungen sind in der Realität aus technischen Gründen nicht aufrechtzuerhalten. Eine wichtige Rolle spielen hierbei die Begriffe der partiellen zeitlichen und der partiellen räumlichen Kohärenz, die ursächlich das Auflösungsvermögen eines Mikroskops begrenzen. Beide Mechanismen bewirken eine zu höheren Frequenzen hin zunehmende selektive Dämpfung der im Bild enthaltenen Fourierkomponenten.

Die in Gleichung (3.15) dargestellte Bildintensität $I(\vec{r})$ läßt sich ebenso wie die Wellenfunktion in Fourierkomponenten zerlegen:

$$I_{\text{Bild}}(\vec{r}) = \sum_{\vec{q}} I(\vec{q}) \exp(2\pi i \vec{q}\vec{r}). \quad (3.19)$$

Die hier dargestellten *Bildfrequenzen* $|\vec{q}|$ dürfen nicht mit den *Raumfrequenzen* $|\vec{g}|$ der Austrittswellenfunktion verwechselt werden, wie anhand des Beispiels in Gleichung (3.18) dargelegt wurde. Die Dämpfung der Fourierkoeffizienten $I(\vec{q})$ wird wegen der begrenzten Kohärenz des einfallenden Strahls für höhere Bildfrequenzen, die ja kleine Objektabstände beschreiben, total, was einem Abschneiden der Fourierreihe in Gleichung (3.19) entspricht.

3.3 Beschreibung durch numerische Bildsimulation

Das Konzept der numerischen Bildsimulation ist eine 1:1-Übersetzung des aktuellen theoretischen Wissens der im letzten Abschnitt beschriebenen physikalischen Vorgänge in eine Programmiersprache. Deshalb ist das Verständnis der Bildsimulation eng verknüpft mit dem Verständnis der zugrundeliegenden Physik, nämlich der dynamischen Elektronenbeugung und der nichtlinearen, partiell kohärenten Optik. Dieses Hintergrundwissen ist sowohl für die sorgfältige experimentelle Arbeit im Grenzbereich des Auflösungsvermögens heutiger Elektronenmikroskope als auch für die Interpretation der gewonnenen Bilder unerlässlich.

Gemäß den Erläuterungen des letzten Abschnitts erfolgt die Bildsimulation in zwei Stufen. Es sind dies:

- Simulation der Elektronenbeugung
- Simulation des Abbildungsprozesses

3.3.1 Simulation der Elektronenbeugung

Ziel der ersten Stufe der Bildsimulation ist die Berechnung der Austrittswellenfunktion oder genauer gesagt ihrer komplexen Fourierkoeffizienten. In diesem Schritt, der der oben beschriebenen Fourierzerlegung der Objektstruktur entspricht, werden die Amplituden und Phasen der einzelnen Teilwellen Ψ_n ermittelt.

Charakteristisch für den heutigen Stand der Bildsimulation ist, daß die Berechnung – abgesehen von der Energie des einfallenden Elektronenstrahls – *vollkommen unabhängig* von den speziellen Eigenschaften des Mikroskops erfolgt. Man geht also von einer ebenen Welle genau definierter Energie aus, die auf das Präparat trifft. Abweichungen von dieser Idealisierung werden entgegen dem zeitlichen physikalischen Ablauf erst im nächsten Schritt einbezogen.

Zur Simulation der Elektronenbeugung stehen zwei grundsätzlich verschiedene Verfahren zur Verfügung, die im nächsten Kapitel erläutert werden:

- Das *Blochwellenverfahren* und
- das *Multislit-Verfahren*

3.3.2 Simulation des Abbildungsprozesses

Im Gegensatz zur Berechnung der Austrittswellenfunktion ist die Simulation der Abbildungsvorgänge ausschließlich durch die physikalischen Eigenschaften des Mikroskops bestimmt. Hier interessiert die Frage, wie die im ersten Schritt berechnete Austrittswellenfunktion durch die in Abschnitt 3.2.2 beschriebenen Vorgänge modifiziert wird. Am Ende dieses Simulationsabschnittes steht die graphische Darstellung der berechneten Bildintensität, die zum optischen Vergleich mit experimentellen Bildern herangezogen wird.

Kapitel 4

Berechnung der Austrittswellenfunktion

4.1 Das Blochwellenverfahren

Im Gegensatz zur kinematischen Beugungstheorie berücksichtigt die dynamische Theorie die Wechselwirkung der einzelnen Teilwellen bei der Durchstrahlung des Objekts. Ausgangspunkt beim Bethe-Bloch-Verfahren [42, 43] ist die relativistisch korrigierte Schrödingergleichung für die Wellenfunktion $\Psi(\vec{r})$ im Kristallpotential $V(\vec{r})$:

$$\nabla^2 \Psi(\vec{r}) + \frac{2m|e|}{\hbar^2} [E^* + V(\vec{r})] \Psi(\vec{r}) = 0 \quad (4.1)$$

wobei m die relativistisch korrigierte Elektronenmasse ist. E^* stellt das relativistisch korrigierte beschleunigende Potential dar, so daß die totale Energie eines einfallenden Elektrons $|e|E^*$ ist:

$$E^* = E \frac{2E_0 + |e|E}{2(E_0 + |e|E)} \quad (4.2)$$

Hierbei sind:

- $E_0 = m_0 c^2$: Ruheenergie des Elektrons
- E : Beschleunigungsspannung des Mikroskops
- e : Ladung des Elektrons

Das periodische Kristallpotential läßt sich mittels der reziproken Gittervektoren als Fourierreihe ausdrücken:

$$V(\vec{r}) = \sum_h V_h \exp(2\pi i \vec{h} \vec{r}) \quad (4.3)$$

Die Koeffizienten V_h sind mit den Strukturamplituden F_h (Gl. (3.14)) durch die Beziehung

$$V_h = \frac{\hbar^2}{2\pi m|e|V_E} F_h^T \quad (4.4)$$

verbunden. Dabei ist V_E das Volumen der Einheitszelle. Die Strukturamplituden werden dabei bezüglich der thermischen Bewegung der Atome, die die Intensität der abgebeugten Strahlen reduziert, durch den *Debye-Waller-Faktor* M korrigiert:

$$F_h^T = F_h \exp(-M) \quad \text{mit} \quad M = Bh^2/4 \quad (4.5)$$

Die temperaturabhängigen Werte für B können in Tabellen nachgeschlagen werden [21].

Für die Wellenfunktion in Gleichung (4.1) wird eine Blochfunktion als Ansatz benutzt:

$$\Psi(\vec{r}) = C(\vec{r}) \exp(2\pi i \vec{k} \vec{r}) \quad (4.6)$$

wobei die Funktion $C(\vec{r})$ die Periodizität des Kristalls besitzt und ebenfalls als Fourierreihe entwickelt werden kann:

$$\Psi(\vec{r}) = \sum_g C_g \exp[2\pi i(\vec{k} + \vec{g}) \cdot \vec{r}] \quad (4.7)$$

Die C_g werden als Blochwellenkoeffizienten bezeichnet. Es ist üblich, die Konstanten zusammenzufassen, indem man die Größe U als modifiziertes Potential benutzt:

$$U_h = \frac{2m|e|}{h^2} V_h \quad (4.8)$$

Die Größe K mit

$$K^2 = \frac{2m|e|}{h^2} (E^* + V_0) \quad (4.9)$$

gibt einen mittleren Wellenvektor der Elektronen im Kristall an. V_0 ist hierbei der Mittelwert des Kristallpotentials. Diese Korrektur gegenüber dem Wellenvektor k der einfallenden Welle ist sehr klein, da typischerweise $E^* \approx 10^5 \text{ V}$, $V_0 \approx 10 \text{ V}$. Indem man die Gleichungen (4.3) und (4.7) bis (4.9) in (4.1) einsetzt, erhält man:

$$\sum_g \left\{ [K^2 - (\vec{k} + \vec{g})^2] C_g + \sum_{h \neq g} U_{g-h} C_h \right\} \exp[2\pi i(\vec{k} + \vec{g}) \cdot \vec{r}] = 0 \quad (4.10)$$

Obige Gleichung muß für alle Orte $\vec{r}$ im Kristall gelten, daher muß jeder Koeffizient der Exponentialfunktion verschwinden. Deshalb erhält man für jeden in die Rechnung einbezogenen Wert von $\vec{g}$ – das heißt für jeden Strahl – eine Gleichung:

$$[K^2 - (\vec{k} + \vec{g})^2] C_g + \sum_{h \neq g} U_{g-h} C_h = 0 \quad (4.11)$$

Dieser Satz von Gleichungen ist exakt, vorausgesetzt man bezieht unendlich viele Strahlen in die Rechnung ein. In der Praxis muß man sich mit einer näherungsweisen Lösung zufrieden geben, da aus Rechenzeitgründen und wegen der begrenzten numerischen Stabilität des nachfolgenden Verfahrens zur Lösung des Gleichungssystems nur eine begrenzte Strahlzahl benutzt werden kann.

Für die Erklärung des Verfahrens eignet sich das einfach nachvollziehbare Beispiel des Zweistrahlfalls. Die Wellenfunktion (4.7) wird beschrieben durch:

$$\Psi(\vec{r}) = C_0 \exp(2\pi i \vec{k} \cdot \vec{r}) + C_g \exp[2\pi i(\vec{k} + \vec{g}) \cdot \vec{r}] \quad (4.12)$$

Mit der Beziehung (4.11) ergeben sich daraus die folgenden Gleichungen:

$$\begin{aligned} [K^2 - (\vec{k} + \vec{g})^2] C_g + U_g C_0 &= 0 \\ (K^2 - k^2) C_0 + U_{-g} C_g &= 0 \end{aligned} \quad (4.13)$$

Diese Gleichungen lassen sich in Matrixform darstellen:

$$\begin{pmatrix} K^2 - k^2 & U_{-g} \\ U_g & K^2 - (\vec{k} + \vec{g})^2 \end{pmatrix} \begin{pmatrix} C_0 \\ C_g \end{pmatrix} = 0 \quad (4.14)$$

Gleichung (4.14) läßt sich lösen, falls die Determinante verschwindet:

$$\begin{vmatrix} K^2 - k^2 & U_{-g} \\ U_g & K^2 - (\vec{k} + \vec{g})^2 \end{vmatrix} = 0 \quad (4.15)$$

Man erhält in diesem Falle eine Gleichung der Ordnung k^4 . Hätte man n Strahlen in die Rechnung einbezogen, ergäbe sich eine Gleichung der Ordnung k^{2n} mit $2n$ Wurzeln. Von diesen sind n positiv und n negativ und entsprechen Vorwärtsstreuung bzw. Rückwärtsstreuung. Die Rückwärtsstreuung spielt jedoch für Elektronen hoher Energie eine untergeordnete Rolle.

Die allgemeine Lösung für die Wellenfunktion $\Psi(\vec{r})$ ist deshalb eine Linearkombination von n Blochwellen mit verschiedenen Wellenvektoren $k^{(j)}$:

$$\Psi(\vec{r}) = \sum_{j=1}^n \alpha^{(j)} \Psi^{(j)}(\vec{r}) = \sum_{j=1}^n \alpha^{(j)} \sum_{g=1}^n C_g^{(j)} \exp [2\pi i (\vec{k}^{(j)} + \vec{g}) \cdot \vec{r}] \quad (4.16)$$

Die Größe $\alpha^{(j)}$ ist die Amplitude der j -ten Blochwelle und wird auch oft *Anregungsstärke* genannt.

Vom physikalischen Standpunkt aus verhält sich der periodische Kristall wie ein Interferometer. Eine einfallende Welle der Energie $|e|E^*$ mit festem Wellenvektor entkoppelt in einen Satz von Blochwellen mit verschiedenen Wellenvektoren $\vec{k}^{(j)}$. Bei der Ausbreitung im Kristall kommen die Blochwellen aufgrund der verschiedenen Wellenlänge außer Phase und interferieren miteinander. Man kann sich diesen Fall am Beispiel eines gekoppelten Pendels klarmachen: Wenn die Schwingungsperioden zweier Pendel nicht zu stark voneinander abweichen, zeigt die gekoppelte Schwingung beider Pendel sogenannte Schwebungen. Diese Schwebungen haben eine viel größere Periode als die Einzelpendel. Die Abweichungen der $k^{(j)}$ vom mittleren Wellenvektor K sowie dem des einfallenden Elektronenstrahls sind aufgrund der minimalen Korrektur der Elektronenenergie durch die V_g ebenfalls nur minimal ($E^*/V_g \approx 10^4$). Die in Gleichung (4.16) dargestellte „Mischung“ verschiedener Blochzustände führt zu Interferenzeffekten, die ähnlich wie im Beispiel mit den gekoppelten Pendeln eine viel größere räumliche Periode haben als die Wellenlänge des einfallenden Elektronenstrahls. Eine Aufspaltung der Wellenfunktion $\Psi(\vec{r})$ (Gl. (4.16)) in die einzelnen Strahlen ergibt dann:

$$\Psi(\vec{r}) = \sum_g \Psi_g = \sum_g A_g \exp (2\pi i \vec{g} \cdot \vec{r}) \quad (4.17)$$

mit den Fourierkoeffizienten

$$A_g(t) = \sum_j \alpha^{(j)} C_g^{(j)} \exp(2\pi i \vec{k}_z^{(j)} t) \quad (4.18)$$

Hierbei gibt t die in Einstrahlungsrichtung durchlaufene Objektdicke an und $\vec{k}_z^{(j)}$ die Komponente des Wellenvektors der j -ten Blochwelle entlang einer bestimmten Zonenachse. Die Gleichungen (4.17) und (4.18) weisen Merkmale auf, die für das Verständnis der späteren Ausführungen wichtig sind:

- Man nimmt formal eine Entkopplung der Wellenpropagation in eine Komponente parallel zur Strahlrichtung und in eine senkrecht zur Strahlrichtung vor. Bei Einstrahlung entlang einer Zonenachse liegen die $\vec{g}$ senkrecht zur Strahlrichtung, weshalb der Exponentialterm in (4.17) eine dazu laterale Ausbreitung beschreibt. (Dies gilt exakt nur, falls – wie im Kapitel *Grundlagen* dargelegt – ausschließlich Strahlen der niedrigsten Lauezone in Betracht gezogen werden, ist aber anschaulich besser vorstellbar.) Die Ausbreitung der Strahlen *entlang* der Einstrahlungsrichtung steckt in den Fourierkoeffizienten A_g .
- Die Summe in (4.18) bewirkt die oben erwähnten Interferenzeffekte, die – wie später deutlicher zu sehen sein wird – auch als Pendellösung bezeichnet werden. Dies bedeutet, daß sich die Interferenzeffekte zwischen den einzelnen Blochwellen in einer mit der Präparatdicke t variierenden Modulation der Fourierkoeffizienten A_g niederschlagen.

Für die Anregungsstärken $\alpha^{(j)}$ erhält man man aus der Kontinuität von Ψ und $\nabla\Psi$ an der Eintrittsfläche des Kristalls im symmetrischen Laue Fall

$$\alpha^{(j)} = C_0^{(j)*} \quad (4.19)$$

Es bleibt nun die Frage zu klären, wie man sich die Blochwellenkoeffizienten $C_g^{(j)}$ beschafft. Mit den Abschätzungen (*Hochenergienäherung*)

$$K \approx k^{(j)}, \quad K \gg g \quad \text{und deshalb} \quad K \approx |\vec{k} + \vec{g}| \quad (4.20)$$

und einer Aufteilung der Blochwellenvektoren $\vec{k}^{(j)}$ in einer zur Laueebene senkrechte Komponente $k_z^{(j)}$ und eine Tangentialkomponente k_t (letztere muß aus den besprochenen Kontinuitätsgründen für alle Blochwellen gleich und gleich der Tangentialkomponente des einfallenden Strahls sein) erhält man:

$$K^2 - k^{(j)2} = K^2 - k_z^{(j)2} - k_t^2 \approx 2K(K - k_z^{(j)}) - k_t^2 \quad (4.21)$$

$$\begin{aligned} K^2 - (\vec{k}^{(j)} + \vec{g})^2 &= (K^2 - k_z^{(j)2}) - (k_t^2 + 2gk_t + g^2) \\ &\approx 2K(K - k_z^{(j)}) - (k_t + g)^2 \end{aligned} \quad (4.22)$$

Setzt man die beiden letzten Gleichungen in (4.14) ein, erhält man:

$$\frac{1}{2K} \begin{pmatrix} -k_t^2 & U_{-g} \\ U_g & -(k_t + g)^2 \end{pmatrix} \begin{pmatrix} C_0^{(j)} \\ C_g^{(j)} \end{pmatrix} = (k_z^{(j)} - K) \begin{pmatrix} C_0^{(j)} \\ C_g^{(j)} \end{pmatrix} \quad (4.23)$$

Aus der direkt einsehbaren Erweiterung des Zweistrahlfalles erhält man folgende Eigenwertgleichung für den Vielstrahlfall:

$$\frac{1}{2K} \begin{pmatrix} -k_t^2 & U_{-g} & U_{-h} & \cdots \\ U_g & -(k_t + g)^2 & U_{g-h} & \cdots \\ U_h & U_{h-g} & -(k_t + h)^2 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} C_0^{(j)} \\ C_g^{(j)} \\ C_h^{(j)} \\ \vdots \end{pmatrix} = (k_z^{(j)} - K) \begin{pmatrix} C_0^{(j)} \\ C_g^{(j)} \\ C_h^{(j)} \\ \vdots \end{pmatrix} \quad (4.24)$$

Diese sogenannte *Säkulargleichung*, die durch Diagonalisieren gelöst werden kann, lautet in Kurzform:

$$M \vec{C}^{(j)} = (k_z^{(j)} - K) \vec{C}^{(j)} \quad (4.25)$$

Wenn n Strahlen in die Rechnung einbezogen werden ist M eine $(n \times n)$ Matrix, es gibt n Eigenwerte $(k_z^{(j)} - K)$ und n Blochwellen, die wiederum jede aus n ebenen Wellen zusammengesetzt sind.

Da $U(\vec{r})$ reell ist, (d.h. $U_g = U_{-g}^*$) ist die Matrix M hermitesch. Deshalb sind die Eigenwerte $(k_z^{(j)} - K)$ reell, die Eigenvektoren $\vec{C}^{(j)}$ sind i.a. jedoch komplex. Falls der Kristall zentrosymmetrisch ist, gilt $U_g = U_{-g}$ und M ist eine reelle symmetrische Matrix. In diesem Falle sind auch die Eigenvektoren $\vec{C}^{(j)}$ reell. Die Matrix C , deren Spalten die Eigenvektoren $\vec{C}^{(j)}$ sind, ist unitär, d.h.

$$\begin{aligned} \sum_g C_g^{(i)} C_g^{(j)*} &= \delta_{ij} \\ \sum_j C_g^{(j)} C_h^{(j)*} &= \delta_{gh} \end{aligned} \quad (4.26)$$

Aus der Bedingung der Kontinuität für die Wellenfunktion an der Eintrittsfläche des Kristalls ergibt sich daher bei einem einfallenden Strahl der Intensität 1 für die Anregungsstärken:

$$\begin{aligned} \sum_j \alpha^{(j)} C_0^{(j)} &= 1 \quad \text{bzw.} \\ C \vec{\alpha} &= \vec{u} \end{aligned} \quad (4.27)$$

Hierbei ist die erste Komponente des Vektors $\vec{u}$ gleich 1, alle anderen Komponenten sind null.

Die gesuchten Fourierkomponenten (Gl. (4.18)) der Austrittswellenfunktion lassen sich in Vektorform schreiben:

$$\vec{a} = \begin{pmatrix} A_0(t) \\ A_g(t) \\ A_h(t) \\ \vdots \end{pmatrix} \quad (4.28)$$

Damit kommt man zu der kompakten Formulierung der Fourierkomponenten $\vec{a}$:

$$\vec{a} = C [\exp(2\pi i k_z^{(j)} t)]_D C^{-1} \vec{u} \quad (4.29)$$

wobei $[\dots]_D$ eine Diagonalmatrix bedeutet. Die Amplituden der abgebeugten Strahlen bei einer bestimmten Objektdicke t sind daher *vollständig* durch die Eigenvektoren und Eigenwerte der Matrix (4.24) definiert.

Einbeziehung von Absorption

Analog zur Lichtoptik, wo Absorption durch die Einführung eines komplexen Brechungsindex beschrieben wird, behandelt man die inelastische Streuung von Elektronen phänomenologisch durch die Einbeziehung eines Imaginärteils im Kristallpotential $V(\vec{r})$.

$$\begin{aligned} V(\vec{r}) &\longrightarrow V(\vec{r}) + iV'(\vec{r}) \\ V_g &\longrightarrow V_g + iV'_g \\ U_g &\longrightarrow U_g + iU'_g \end{aligned} \quad (4.30)$$

Die wichtigsten in Betracht kommenden inelastischen Streuprozesse sind:

- Plasmonenanregung
- Elektron – Elektron Streuung
- Phononenstreuung

Diese Prozesse vermindern die zur Abbildung beitragende Zahl der Elektronen durch zwei Effekte:

- Elektronen werden unter so großen Winkeln abgelenkt, daß sie die Objektiveblende nicht mehr passieren können.
- Elektronen erfahren unterschiedliche Energieverluste, so daß die ihnen zugeordneten Wellen untereinander nicht mehr kohärent sind.

Die Einführung komplexer Potentiale in der zu Anfang des Kapitels stehenden Schrödingergleichung führt zu komplexen Blochwellenvektoren $\vec{k}^{(j)}$:

$$\vec{k}^{(j)} \longrightarrow \vec{k}^{(j)} + i\vec{q}^{(j)} \quad (4.31)$$

Die Wellenfunktion in Gleichung (4.16) lautet damit:

$$\Psi(\vec{r}) = \sum_j \alpha^{(j)} \sum_g C_g^{(j)} \exp [2\pi i(\vec{k}^{(j)} + \vec{g}) \cdot \vec{r}] \exp (-2\pi \vec{q}^{(j)} \cdot \vec{r}) \quad (4.32)$$

Dies bedeutet, daß jede Blochwelle bei der Durchquerung des Kristalls exponentiell gedämpft wird. Ohne Absorption sind die Anregungsstärken $\alpha^{(j)}$ über die Kristalldicke konstant. Der Effekt der Absorption läßt sich formal auf die Anregungsstärken überwälzen, die damit dickenabhängig werden:

$$\alpha^{(j)}(t) = \alpha^{(j)} \exp (-2\pi |\vec{q}^{(j)}| t) \quad (4.33)$$

Die Tatsache, daß verschiedene Blochwellen jeweils unterschiedlich stark gedämpft werden und sich damit das „Mischungsverhältnis“ der Blochwellen in der Wellenfunktion $\Psi(\vec{r})$ laufend mit der Präparatdicke ändert, wird *anormale Absorption* genannt. Dieses Verhalten läßt sich qualitativ folgendermaßen erklären:

Blochwellen lassen sich als *laufende Wellen in Einfallsrichtung* des Elektronenstrahls auffassen, jedoch als *stehende Wellen in einer Ebene senkrecht dazu*. Bestimmte Blochwellen sind in der Nähe der Atome lokalisiert, andere haben größere Aufenthaltswahrscheinlichkeit in Regionen zwischen den Atomen. Dieses Verhalten kann man an den Beträgen der Wellenvektoren $\vec{k}^{(j)}$ ablesen, nach denen sich die Blochwellen der Reihe nach ordnen lassen. Blochwellen, die in Atomnähe lokalisiert sind, haben geringe potentielle Energie und somit aus Gründen der Energieerhaltung hohe kinetische Energie und damit größeres $|\vec{k}^{(j)}|$ als Blochwellen die dazwischen lokalisiert sind. Die physikalische Ursache für anormale Absorption ist nun, daß der periodische Imaginärteil des Potentials $V'(\vec{r})$ Maxima an den Plätzen innerhalb der Einheitszelle hat, wo die inelastische Streuung am stärksten ist. Eine Betrachtung der oben erwähnten Streumechanismen zeigt, daß die Phononenstreuung stark an den Atomplätzen lokalisiert ist, Elektron - Elektron Streuung relativ schwach an den Atomplätzen lokalisiert ist und Plasmonenanregung kaum lokalisiert ist. Aus diesem Grunde hat $V'(\vec{r})$ seine Maxima ebenfalls in Atomnähe. Blochwellen, die selbst stark an den Atomplätzen lokalisiert sind, die also großes $|\vec{k}^{(j)}|$ haben, werden am stärksten gedämpft.

Aufgrund der qualitativen Behandlung der Absorption läßt sich der Imaginäranteil des Gitterpotentials

$$V'(\vec{r}) = \kappa V(\vec{r}) \quad (4.34)$$

schwer berechnen. In der Literatur (z.B. [43]) wird die Größenordnung $\kappa \leq 0.1$ als geeigneter Wert zur Beschreibung der Absorption angegeben.

Die Dämpfungskoeffizienten $q^{(j)}$ lassen sich unter der Annahme, daß die Koeffizienten C_g unverändert bleiben, durch einfaches Einsetzen der komplexen Potentiale in (4.11) ebenfalls in Matrixform darstellen:

$$q^{(j)} = \frac{1}{2K} (C_0^{(j)*} C_g^{(j)*} \dots) \begin{pmatrix} U'_0 & U'_{-g} & \dots \\ U'_g & U'_0 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} C_0^{(j)} \\ C_g^{(j)} \\ \vdots \end{pmatrix} \quad (4.35)$$

4.2 Das Multislice-Verfahren

Im Multislice-Verfahren [44] wird die Wechselwirkung der einfallenden Elektronenwelle mit dem periodischen Gitterpotential $V(\vec{r})$ als Abfolge von Phasenverschiebungen, die die Wellenfunktion $\Psi(\vec{r})$ abrupt beim Durchlaufen hintereinander liegender Beugungsebenen erfährt, beschrieben. Ein „slice“ besteht dabei (Abb. 4.1) aus zwei Komponenten:

- Beugungsebene

In den senkrecht zur Einfallsrichtung liegenden infinitesimal dünnen Beugungsebenen wird formal der Phasenverschiebung Rechnung getragen, die die Wellenfunktion $\Psi(\vec{r})$ beim Durchqueren einer Kristallschicht der Dicke Δz aufgrund der Einwirkung des Kristallpotentials erfahren hat. Die streuenden Atompotentiale innerhalb der Schicht Δz werden sozusagen auf eine Ebene „komprimiert“ und müssen deshalb geeignet modifiziert werden. Dieser Vorgang wird durch die sogenannte *Phasengitterfunktion* beschrieben.

- Leerer Raum der Dicke Δz

Durch die Beschreibung der Ausbreitung der Wellenfunktion $\Psi(\vec{r})$ zwischen zwei Beugungsebenen wird der Dreidimensionalität des Kristalls Rechnung getragen, die durch bloßes Aneinanderreihen der infinitesimal dünnen Beugungsebenen verloren ginge. Dies geschieht mit Hilfe der *Propagatorfunktion*.

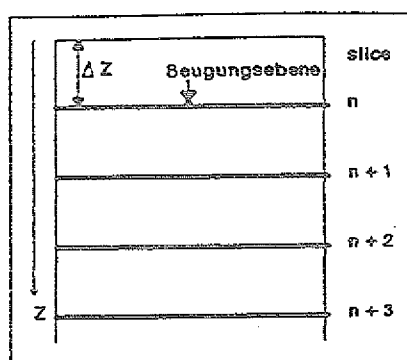


Abbildung 4.1: Darstellung der Abfolge von vier „slices“ entlang der Richtung des einfallenden Elektronenstrahls. Ein „slice“ hat die Dicke Δz .

Phasengitterfunktion

Die einfallende Elektronenwelle sei gegeben durch:

$$\Psi(\vec{r}) = \Psi_0 \exp(2\pi i \vec{k}_0 \vec{r}) \quad (4.36)$$

Daraus erhält man für den Wellenvektor $\vec{k}_0$ im Vakuum nach relativistischer Korrektur des beschleunigenden Potentials E und der Masse m :

$$|\vec{k}_0| = \left(\frac{2m|e|E^*}{h^2} \right)^{1/2} \quad (4.37)$$

Innerhalb des Kristalls ändert sich der Wellenvektor zu

$$|\vec{k}_i(\vec{r})| = \left(\frac{2m|e|(E^* + V(\vec{r}))}{\hbar^2} \right)^{1/2} \quad (4.38)$$

Die Wurzel in der vorangehenden Gleichung läßt sich entwickeln ($V \ll E^*$) und der Wellenvektor im Kristall lautet mit dieser Näherung:

$$|\vec{k}_i(\vec{r})| \approx |\vec{k}_0| \left(1 + \frac{V(\vec{r})}{2E^*} \right) = |\vec{k}_0| + \frac{m|e|V(\vec{r})}{\hbar^2 |\vec{k}_0|} \quad (4.39)$$

Nach dem Durchqueren einer Schicht der Dicke dz ist die Phase der Welle gegenüber der Phase einer im Vakuum laufenden Welle verschoben. Die im Kristall laufende Welle läßt sich damit folgendermaßen darstellen:

$$\Psi(x, y, z + dz) = \underbrace{\Psi_0 \exp(2\pi i \vec{k}_0 \vec{r})}_{\Psi(x, y, z)} \exp \left[\frac{2\pi i m |e| V(\vec{r}) dz}{\hbar^2 |\vec{k}_0|} \right] \quad (4.40)$$

Durch Integration über die in Abb.4.1 dargestellte Strecke Δz , die sich typischerweise über einen Gitterebenenabstand erstreckt, erhält man:

$$\Psi(x, y, z + \Delta z) = \Psi(x, y, z) \cdot Q \quad (4.41)$$

mit der *Phasengitterfunktion*

$$Q = \exp \left[\frac{2\pi i m |e|}{\hbar^2 |\vec{k}_0|} \int_0^{\Delta z} V(x, y, z) dz \right] \quad (4.42)$$

Die Phasenverschiebung der Wellenfunktion $\Psi(\vec{r})$ aufgrund der streuenden Atom-potentiale in einem „slice“ läßt sich durch simples Multiplizieren mit der Phasengitterfunktion beschreiben. Durch Berechnung des über die Schichtdicke konstanten sogenannten *projizierten Potentials* $V_p(x, y)$, mit

$$V_p(x, y) = \frac{1}{\Delta z} \int_0^{\Delta z} V(x, y, z) dz \quad (4.43)$$

aus den Fourierkoeffizienten V_{hkl} [45], kann man die Phasengitterfunktion folgendermaßen darstellen:

$$Q(x, y) = \exp \left[\frac{2\pi i m |e|}{\hbar^2 |\vec{k}_0|} V_p(x, y) \Delta z \right] \quad (4.44)$$

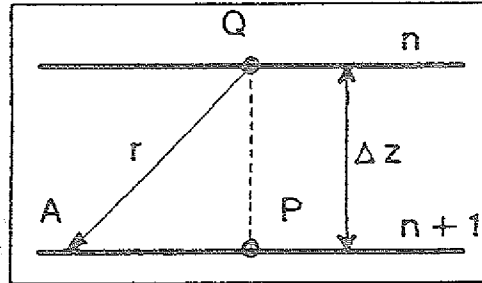


Abbildung 4.2: Darstellung zur Geometrie bei der Herleitung der Propagatorfunktion

Propagatorfunktion

Die Ausbreitung der Elektronenwelle zwischen den „slices“ n und $n+1$ wird durch den *Fresnelpropagator* beschrieben: Jeder Punkt der Beugungsebene n (Abb. 4.2) ist nach dem Huygens-Fresnelschen Prinzip Quelle einer kohärenten sphärischen Welle, deren Einzelbeiträge sich in der Ebene $n+1$ aufsummieren. Dabei ist der Abstand r der Quellpunkte Q zu den Punkten der Ebene $n+1$, an denen die Beiträge aufsummiert werden müssen, je nach Lage der Quellpunkte verschieden. Das bedeutet, daß die sphärischen Wellen je nach Emissionsort in der Ebene n mit verschiedener Phase an einem bestimmten Punkt (z.B. A in Abb. 4.2) ankommen. Eine vom Emissionsort (X, Y) in der Beugungsebene n ausgehende Kugelwelle läßt sich darstellen durch:

$$\Psi \sim \Psi_n(X, Y) \exp(2\pi i k r) \quad (4.45)$$

Aufintegrieren über alle Quellpunkte Q in der Ebene n ergibt am Punkt (x, y) in der Ebene $n+1$:

$$\Psi_{n+1}(x, y) \sim \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Psi_n(X, Y) \exp(2\pi i k r) dX dY \quad (4.46)$$

Unter parabolischer Näherung der von den Quellpunkten Q in der Ebene n ausgehenden sphärischen Wellenfronten wird aus der vorangehenden Gleichung:

$$\Psi_{n+1}(x, y) = \frac{i}{\lambda \Delta z} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \Psi_n(X, Y) \exp \left\{ \frac{\pi i k}{\Delta z} [(x - X)^2 + (y - Y)^2] \right\} dX dY \quad (4.47)$$

Hierbei wurde der allen von den Quellpunkten Q ausgehenden Strahlen gemeinsame Phasenfaktor $\exp(2\pi i \Delta z / \lambda)$ weggelassen. Die Normierungskonstante vor den Integralen folgt aus der Intensitätserhaltung. Dieser Ausdruck ist eine Faltung der Wellenfunktion in der Ebene n mit dem Fresnelpropagator P d.h.

$$\Psi_{n+1}(x, y) = \Psi_n(x, y) * P(x, y) \quad (4.48)$$

mit der Propagatorfunktion

$$P(x, y) = \frac{i}{\lambda \Delta z} \exp \left(\frac{\pi i k (x^2 + y^2)}{\Delta z} \right) \quad (4.49)$$

Der Multislice-Zyklus

Die Wellenfunktion unmittelbar unterhalb der Beugungsebene $n + 1$ lautet nach dem dort erfolgten Phasensprung:

$$\Psi_{n+1}(x, y) = [\Psi_n(x, y) * P_n(x, y)] \cdot Q_{n+1}(x, y) \quad (4.50)$$

Durch sequentielle Anwendung dieser rekursiven Formulierung der Wellenfunktion läßt sich die Austrittswellenfunktion bei einer bestimmten Dicke t , die ein Vielfaches von Δz sein muß, berechnen. Die sehr aufwendige Faltungsoperation führt zu Rechenzeiten, die mit dem Quadrat der Anzahl N der berechneten Punkte (x, y) ansteigt. (In Gleichung (4.47) muß für jeden der N Punkte (x, y) über N Punkte (X, Y) summiert werden.) Darüberhinaus wird eine Formulierung der Austrittswellenfunktion im Fourierraum benötigt, um den Einfluß der Objektivlinse auf die Bildentstehung beschreiben zu können. Mit dem bekannten Faltungstheorem ($\mathcal{F}$ bedeutet Fouriertransformation)

$$\mathcal{F}[\Psi_n(x, y) * P_n(x, y)] = \mathcal{F}[\Psi_n(x, y)] \mathcal{F}[P_n(x, y)] = \Psi_n(\vec{k}) P_n(\vec{k}) \quad (4.51)$$

wird die Faltung durch eine Multiplikation im Fourierraum ersetzt. Damit wird aus Gleichung (4.50)

$$\Psi_{n+1}(\vec{k}) = \mathcal{F} \left\{ \mathcal{F}^{-1}[\Psi_n(\vec{k}) P(\vec{k})] Q_{n+1}(x, y) \right\} \quad (4.52)$$

Der immense Vorteil dieser Manipulation liegt darin, daß die im Realraum aufwendige Faltung umgangen werden kann und durch eine Multiplikation im Fourierraum ersetzt wird. Hierbei kommt die Methode der *schnellen Fouriertransformation* (engl. Fast Fourier Transform, FFT) nach dem sogenannten Cooley-Tukey-Algorithmus zum Einsatz. Obwohl beim Multislice-Verfahren der physikalische Hintergrund der dynamischen Elektronenbeugung nicht mehr nachvollziehbar ist, nennt man die N bei der FFT benutzten Stützstellen ebenfalls Strahlen. Je größer die Ausdehnung der Einheitszelle in der xy -Ebene ist, desto mehr Stützstellen werden benötigt, desto mehr Strahlen müssen in die Rechnung einbezogen werden, um eine ausreichende Auflösung im Ortsraum zu erhalten. Geht man von N benötigten Strahlen aus, ist die bei der FFT benötigte Rechenzeit proportional zu $N \log_2 N$, während beim Blochwellenverfahren die Rechenzeit proportional zu N^2 ist.

Aus diesem Grunde eignet sich das Multislice-Verfahren für die Berechnung der Austrittswellenfunktion „großer“ Einheitszellen, sogenannter *Superzellen*, die man zur Darstellung von Abweichungen der regelmäßigen Kristallstruktur braucht. Sowohl im Multislice- als auch im Blochwellenverfahren werden als Ausgangspunkt der Berechnung die Potentialkoeffizienten V_{hkl} gebraucht. Voraussetzung für die Fourierentwicklung des Gitterpotentials $V(\vec{r})$ in Fourierkomponenten ist jedoch die Periodizität dieser Funktion, die bei Abweichungen von der regelmäßigen Struktur nicht mehr gegeben ist. Durch periodische Fortsetzung der Atomanordnung einer Superzelle in x - und y -Richtung (siehe Abb. 4.3) läßt sich die geforderte Periodizität jedoch wieder herstellen. Typische Superzellen enthalten mehrere hundert Atome, der Aufwand zur Berechnung der Wellenfunktion ist daher nur mehr mit dem Multislice-Verfahren in annehmbaren Rechenzeiten zu bewältigen.

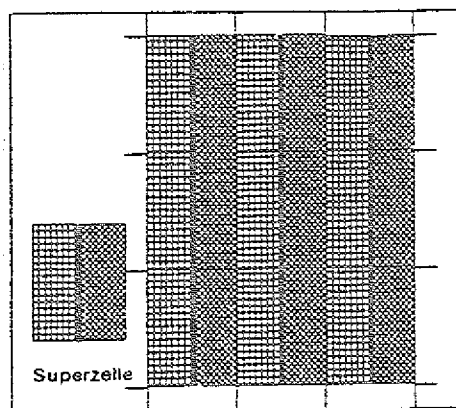


Abbildung 4.3: Periodische Fortsetzung von Superzellen

4.3 Einsatz des EMS-Programmpakets

Zur Berechnung der Austrittswellenfunktion periodischer Objekte wurde das EMS-Programmpaket von P. Stadelmann eingesetzt [46]. Das EMS-Paket besteht aus einer Vielzahl von Programmen, deren Schwerpunkt die Lösung von Aufgabenstellungen aus dem Bereich der Elektronenmikroskopie ist. Angefangen von der perspektivischen Darstellung von Kristallstrukturen über die Berechnung von Beugungsbildern bis hin zur konkreten Durchführung der oben besprochenen Methoden bietet das Paket eine Fülle von interaktiven Programmen, deren vollständige Aufzählung an dieser Stelle zu weit führen würde. Ein Überblick über die für die Simulation hochauflöster Abbildungen wichtigen Programme ist in Abb. 4.4 auf Seite 41 zu sehen. Die zur Simulation periodischer Wellenfunktionen eingesetzten Programme werden im folgenden erläutert.

Ausgangspunkt für die Berechnung der Wellenfunktion regelmäßiger Strukturen ist das Programm BU1, das die Fourierkoeffizienten V_{hkl} des Kristallpotentials berechnet. Das Programm besitzt einen Raumgruppengenerator, der nach Angabe der Raumgruppe (z.B. Nr. 87 für das Hermann-Mauguin Symbol $I4/m$ der $D1_a$ Struktur) sowie der Gitterkonstanten und der Atomsorten den Kristall automatisch „aufbaut“.

Für das Blochwellenverfahren existieren die drei Programmversionen BZ1, BZ2 und BZ3. BZ1 bezieht nur Strahlen der nullten Lauezone ein und eignet sich nur für zentrosymmetrische Kristalle. Die höheren BZ-Versionen arbeiten ohne diese Einschränkungen. Die Einbeziehung höherer Lauezonen spielt bei Kristallen, die eine große Ausdehnung der Einheitszelle entlang der Zonenachse haben, eine Rolle, da die entsprechenden reziproken Gitterebenen im Fourierraum dicht gedrängt sind [45]. Die Maximalzahl der einbeziehbaren Strahlen ist 500. Eine größere Strahlzahl ist wegen der bei der Matrixdiagonalisierung entstehenden numerischen Probleme, die aufgrund der endlichen Genauigkeit der Zahlendarstellung entstehen, problematisch.

Die Programme PG1 und PG2 dienen zur Berechnung des Fresnelpropagators und der Phasengitterfunktion. Beide Programme erlauben eine Unterteilung der Einheitszelle in sogenannte „slices“ entlang der Richtung des einfallenden Strahls. Dies bedeutet, daß mehrere Multislice-Iterationen pro Einheitszelle gemacht wer-

den, was die Genauigkeit der Rechnung erhöht. Beim Programm PG1 wird das projizierte Potential unverändert aus der Integration über die gesamte Einheitszelle übernommen. Dabei wird das projizierte Potential einfach durch die Anzahl der „slices“ geteilt. Beim Programm PG2 wird hingegen das projizierte Potential für jedes „subslice“ innerhalb der Einheitszelle neu berechnet, d.h. man erhält mehrere *verschiedene* Phasengitterfunktionen entlang der Richtung des einfallenden Strahls [45]. Letztenendes bedeutet die immer häufigere Berechnung des projizierten Potentials eine immer genauer werdende Beschreibung des realen Potentials. Ebenso werden die bei der parabolischen Näherung der Propagatorfunktion gemachten Fehler durch Unterteilung der Einheitszelle in mehrere „subslices“ minimiert. (Der Multislice-Formalismus läßt sich aus der Schrödingergleichung ableiten [47] und es läßt sich zeigen, daß die Ergebnisse bei einer gegen null gehenden „slice“-Dicke mit denen des Blochwellenverfahrens identisch sind [48].)

Das Programm MS1 beinhaltet den oben beschriebenen Multislice-Algorithmus zur Berechnung der Austrittswellenfunktion. Für die schnelle Fouriertransformation können in der bestehenden Version maximal 1024×1024 Strahlen einbezogen werden.

4.4 Ergebnisse

Sowohl die Blochwellenprogramme als auch das Multislice-Programm MS1 erlauben das Anlegen einer Datei, in der die komplexen Fourierkoeffizienten A_g der Wellenfunktion verzeichnet sind. Diese Koeffizienten lassen sich aufgespalten nach Amplitude und Phase als Funktion der Präparatdicke darstellen:

$$A_g(t) = |A_g(t)| \exp \{i\Phi(t)\} \quad (4.53)$$

In Abb. 4.5 auf Seite 44 ist als Beispiel das Ergebnis der Simulation für die [001] Zonenachse von Ni_4Mo für einige ausgewählte Strahlen dargestellt. Die zugehörigen Simulationsparameter sind:

- Beschleunigungsspannung: 400 kV
- Strahlzahl:
 - Blochwellenverfahren: 205
 - Multislice-Verfahren: 64×64
- Debye-Waller-Faktor B: 0.0035 nm^2 für Ni, 0.0025 nm^2 für Mo
- Absorptionskonstante κ : 0.05

Beim Blochwellenverfahren wurde die Strahlzahl ausgehend von 50 jeweils um ca. 50 erhöht. Es zeigte sich ausreichende Konvergenz der Rechnung bei Einbeziehung von 150 Strahlen, da die Ergebnisse unter Berücksichtigung von 205 Strahlen sich kaum noch änderten. Zusätzlich wurde als Vergleich die Berechnung mittels der

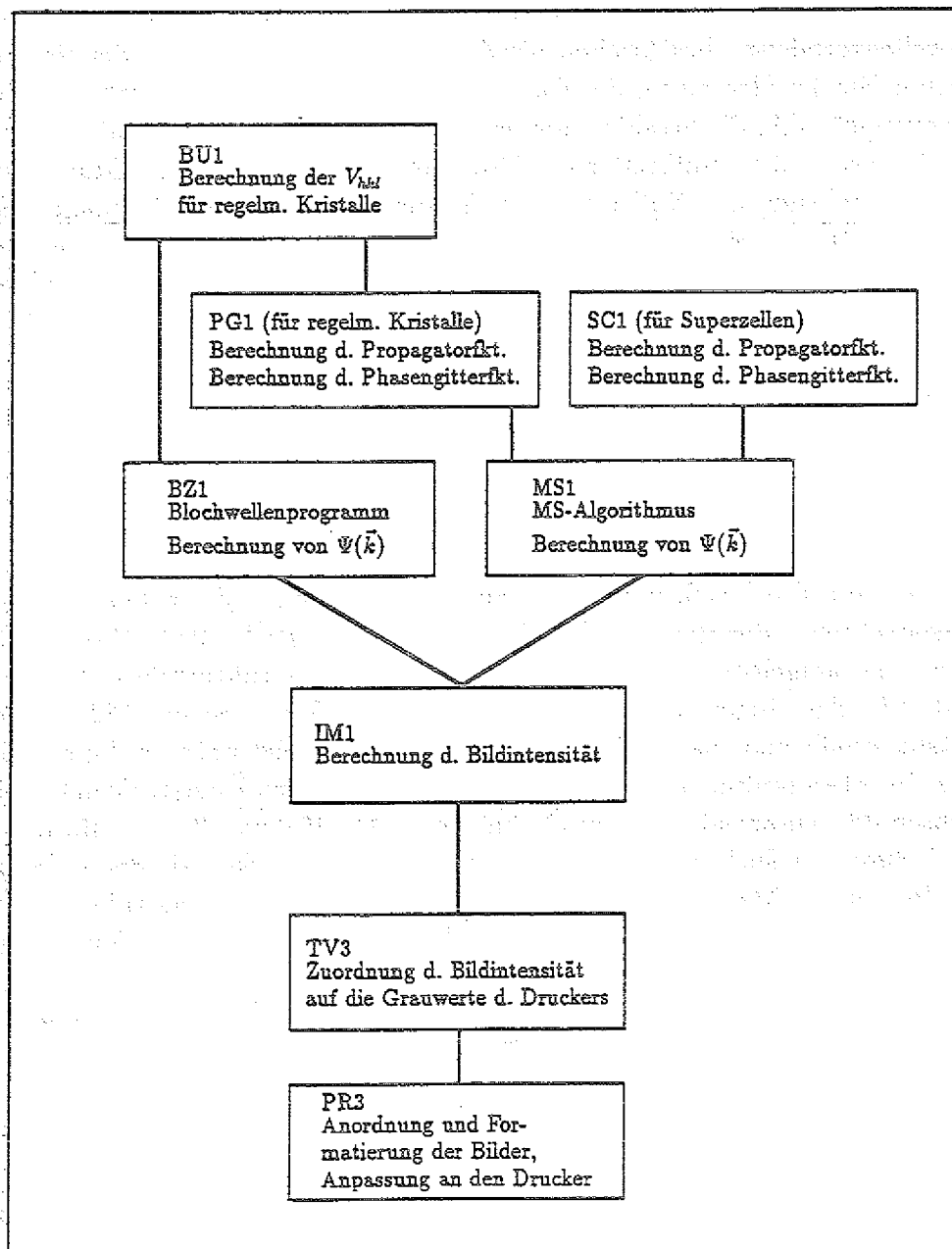


Abbildung 4.4: Blockschema der Simulationsprogramme des EMS-Pakets. Höhere Versionen, die im Text beschrieben werden, sind nicht eingezeichnet.

Multislice-Methode herangezogen. Hierbei wurde, wie auch in späteren Berechnungen, eine Aufteilung der Einheitszelle in zwei „subslices“ vorgenommen. Die aus den beiden Verfahren gewonnenen Ergebnisse waren nahezu identisch.

In Abb. 4.5 sieht man, daß die Amplituden des Nullstrahls sowie der Matrixreflexe (130) und (420) deutlich ein mit der Präparatdicke oszillierendes Verhalten aufweisen, das als „Pendellösung“ bezeichnet wird. Dies kommt – wie im Abschnitt „Blochwellenverfahren“ beschrieben wurde – durch die Interferenz der im Kristall angeregten Blochwellen zustande. Die Oszillation der einzelnen Strahlen ist dynamisch gekoppelt, d.h. die Strahlen können nicht unabhängig voneinander betrachtet werden. So weisen die Amplituden der Matrixreflexe immer dann ein Maximum auf, wenn die Amplitude des Nullstrahls ein Minimum hat. Dieses Phänomen ist aus Gründen der Intensitätserhaltung leicht verständlich. Dieses Verhalten ist an den Überstrukturreflexen nicht mehr so deutlich zu sehen, da diese eine viel größere *Extinktionslänge* aufweisen. Unter der Extinktionslänge ξ_g versteht man den Abstand zweier Amplitudenmaxima eines Strahls. Für eine qualitative Erklärung dieses Verhaltens läßt sich die kinematische Theorie kurz heranziehen. Danach ergibt sich für die Extinktionslänge:

$$\xi_g \sim \frac{1}{V_g} \quad (4.54)$$

Aus der vom Blochwellenprogramm angelegten Datei, in der neben Eigenwerten und Eigenvektoren der dynamischen Beugungsmatrix auch deren Matrixelemente V_{hkl} bzw. U_{hkl} aufgelistet sind, erhält man für den Überstrukturreflex (110): $V_{110} = 3.6$ Volt. Für das Gitterpotential V_{130} , das mit dem Matrixreflex (130) in Verbindung steht, erhält man einen Wert von 20.9 V. Daraus folgt nach der kinematischen Theorie eine etwa fünfmal größere Extinktionslänge für den Überstrukturstrahl. Das Anwachsen der Amplitude ist – im Gegensatz zu den Matrixreflexen – für die Überstrukturreflexe im Dickenbereich unter 100 nm sogar stärker als der gegenläufige Absorptionseffekt. Man sieht an diesem Beispiel, daß die Überstrukturreflexe zu größeren Dicken hin relativ zu den Matrixreflexen zunehmend an Bedeutung gewinnen.

Die gepunkteten Linien in Abb. 4.5 zeigen die Phasen der gebeugten Strahlen relativ zu der des Nullstrahls. Die Beschriftung der y-Achse erfolgt dabei nach folgendem Schema:

$$\begin{aligned} 0 &\equiv -\pi \\ 1 &\equiv +\pi \end{aligned} \quad (4.55)$$

Es ist deutlich zu sehen, daß die Phase der gebeugten Strahlen beim Eintritt in den Kristall bei $\pi/2$ ihren Ausgang nimmt. Dies läßt sich an einem einfachen eindimensionalen Beispiel verdeutlichen: Die einfallende Welle werde mit $\Psi = \Psi_0 \sin(2\pi kz)$ beschrieben. Die gebeugte Welle Ψ_s erfährt nach Durchqueren einer Strecke t im Kristall eine Phasenverschiebung φ :

$$\Psi_s = \Psi_0 \sin(2\pi kz + \varphi) = \Psi_0 \sin(2\pi kz) \cos \varphi + \Psi_0 \cos(2\pi kz) \sin \varphi \quad (4.56)$$

Für *kleine* t gilt dann: $\varphi \ll 1$ und damit

$$\sin \varphi \simeq \varphi \quad \text{und} \quad \cos \varphi \simeq 1 \quad (4.57)$$

$$\Psi_s = \Psi_0 \sin(2\pi kz) + \Psi_0 \varphi \sin(2\pi kz + \pi/2) \quad (4.58)$$

Dieses nur für *kleine* Präparatdicken geltende Beispiel demonstriert zugleich den Begriff des *schwachen Phasenobjekts*. „Schwach“ bezieht sich hierbei auf die Amplitude des abgelenkten Strahls, die wegen $\varphi \ll 1$ im Vergleich zum Nullstrahl klein ist. Wie in Abbildung 4.5 zu sehen ist, gilt die Näherung durch das schwache Phasenobjekt im Falle von Ni_4Mo schon bei Präparatdicken von ca. 5 nm nicht mehr, da sowohl die Amplituden der abgelenkten Strahlen in der Größenordnung des Nullstrahls liegen, als auch die Phasen erheblich vom Wert $\pi/2$ abweichen.

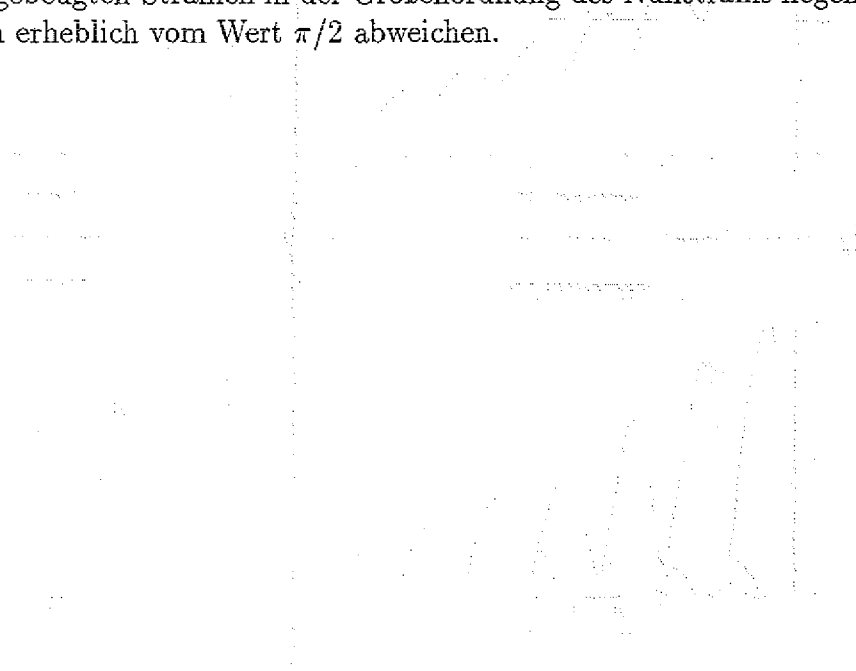


Abbildung 4.5: Vergleich von Amplituden- und Phasenverläufen für ein schwaches Phasenobjekt. Die linke Seite zeigt Amplitudenverläufe für verschiedene Präparatdicken t (0, 5, 10, 20 nm) in Abhängigkeit von der Position z . Die rechte Seite zeigt die entsprechenden Phasenverläufe. Die Amplituden sind in Einheiten von $\Psi_0/2$ skaliert, die Phasen in Einheiten von $\pi/2$. Die Kurven zeigen, dass für kleine t die Amplituden sehr klein sind und die Phase nahe bei $\pi/2$ liegt, während für größere t die Amplituden zunehmen und die Phase davon abweicht.

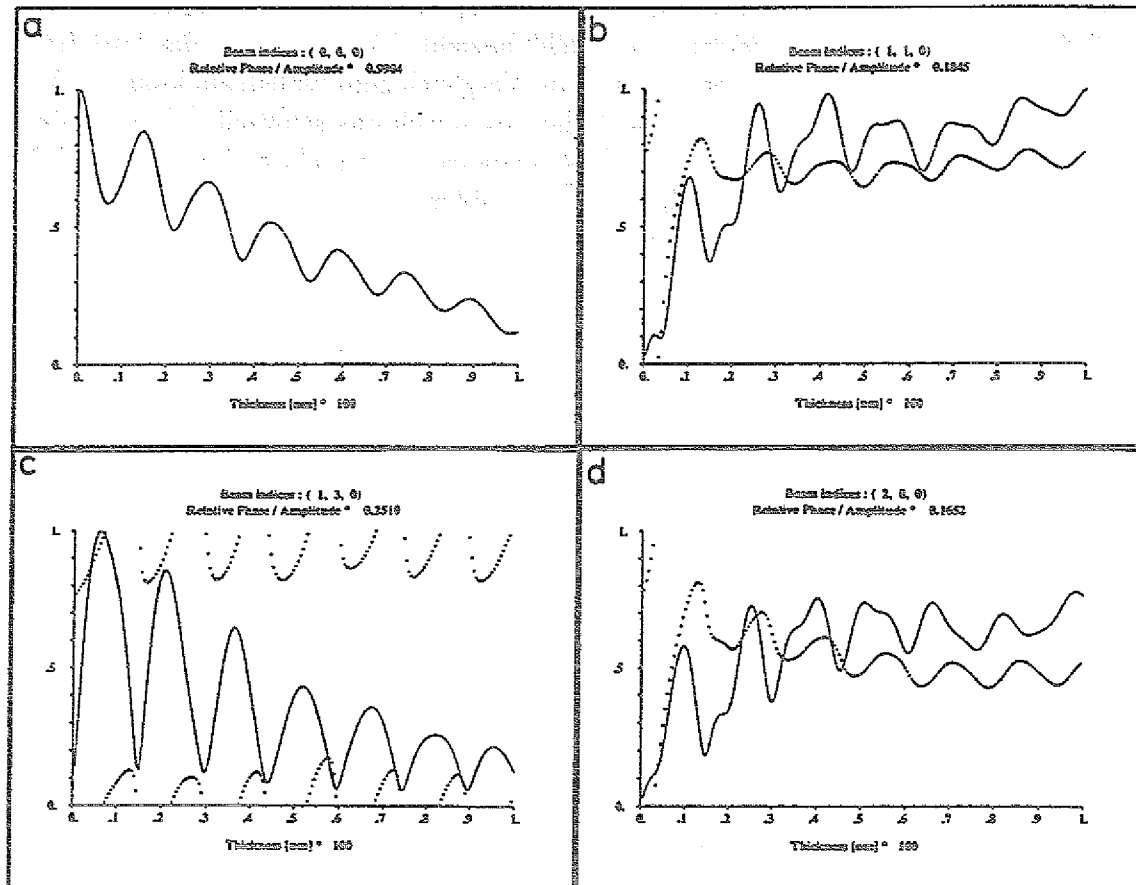


Abbildung 4.5: Amplituden-Phasen Diagramme für einige ausgewählte Strahlen. Die Darstellungen beziehen sich auf die $[001]$ Zonenachse von Ni_4Mo . Durchgezogene Linien: Amplituden. Gepunktete Linien: Phasen. Die Amplituden sind so normiert, daß der einfallende Strahl die Intensität 1 hat. Die Amplituden sind in relativen Einheiten aufgetragen, der jeweilige Maßstabsfaktor ist in den Bildern rechts oben vermerkt. Bei den Phasen entspricht 1 dem Wert $+\pi$ und 0 dem Wert $-\pi$. a: Nullstrahl, b: (110) -, c: (130) -, d: (200) -Strahl.

Kapitel 5

Kontrastübertragung

Im Rahmen dieser Arbeit wurden zwei neuartige Verfahren entwickelt, in deren Mittelpunkt die Auswertung des elektronenoptischen Bildkontrasts steht. In beiden Verfahren, die in den Kapiteln 7 und 8 vorgestellt werden, spielt die Simulation der elektronenoptischen Abbildung eine zentrale Rolle. Zu diesem Zwecke wurden umfangreiche Simulationsprogramme erstellt, die in ihrer numerischen Effizienz für die jeweilige Anwendung maßgeschneidert sind. In diesem Kapitel soll jedoch nicht die konkrete Implementation dieser Programme im Vordergrund stehen, sondern es werden die zum Verständnis der elektronenoptischen Abbildung notwendigen physikalischen Grundlagen erörtert. In den folgenden Ausführungen wird versucht, den nicht ganz trivialen Komplex der partiell kohärenten, nichtlinearen Abbildung Stufe um Stufe verständlich zu erläutern.

5.1 Kontrast des schwachen Phasenobjekts

5.1.1 Kontrast bei kohärenter Einstrahlung

Die in der Elektronenmikroskopie benutzten elektromagnetischen Objektivlinsen bewirken neben einer Vergrößerung des ersten Zwischenbildes durch zwei verschiedene Mechanismen eine Phasendrehung der komplexen Amplituden der Teilwellen der Austrittswellenfunktion, die in hohem Maße von der Raumfrequenz dieser Wellen abhängt. Dies bedeutet, daß die Phasen der einzelnen Strahlen, die im vorangehenden Kapitel berechnet wurden, durch die Objektivlinse regelrecht „durchmischt“ werden. Dieser Effekt ist zum einen durch die nicht vermeidbare *sphärische Aberration* der Objektivlinse konstruktiv vorgegeben, kann aber auch willentlich durch das sogenannte *Defokussieren* der Linse herbeigeführt werden.

Elektromagnetische Linsen lassen sich analog zu lichtoptischen Linsen durch die Gaußsche Linsengleichung beschreiben:

$$\frac{1}{f} = \frac{1}{a} + \frac{1}{b} \quad (5.1)$$

Hierbei bedeutet f die Brennweite der Linse, a die Gegenstandsweite und b die Bildweite. Durch die sphärische Aberration werden nun Strahlen mit verschiedenem

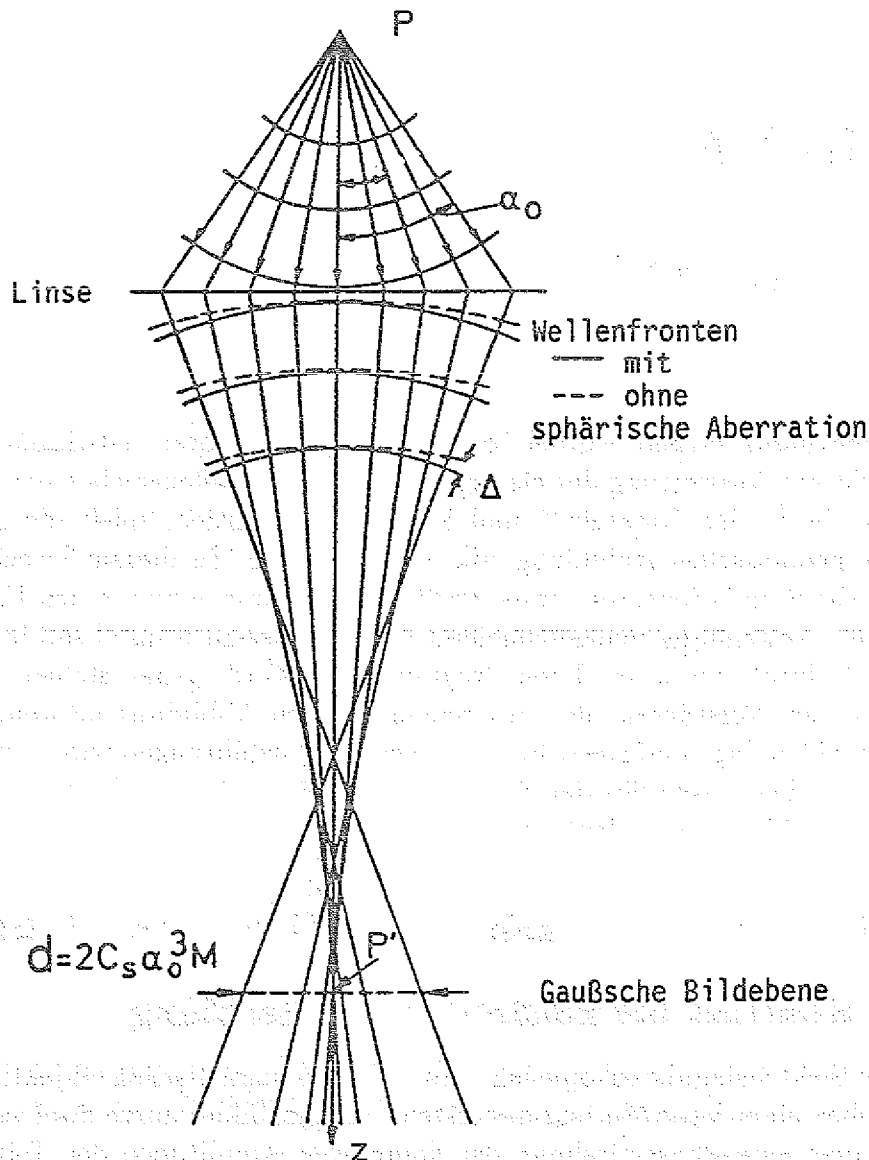


Abbildung 5.1: Elektronentrajektorien und Wellenfronten für eine Linse, die sphärische Aberration aufweist.

Einfallswinkel α_0 zur Objektivlinse (siehe Abb. 5.1) von dieser verschieden stark abgelenkt. Strahlen, die unter größeren Winkeln einfallen, werden stärker abgelenkt als achsnahe Strahlen. Im Wellenbild bedeutet dies, daß die Wellenfronten nach außen hin stärker gekrümmt sind und von der sphärischen Form, die in der Gaußschen Linsengleichung vorausgesetzt wird, abweichen. Dies hat zur Folge, daß Strahlen mit großem α_0 , die von einem Punkt P ausgehen, *vor* der Gaußschen Bildebene gebündelt werden. In der Gaußschen Bildebene sind diese Strahlen bereits wieder divergent. Bezieht man nun alle vom Punkt P ausgehenden Strahlen ein, so erzeugen diese in der Gaußschen Bildebene keinen Punkt sondern eine Scheibe als Bild. Der auf die Objektebene zurückprojizierte Radius dieser Scheibe wird durch die sphärische Aberrationskonstante C_s beschrieben. Für die Bildebene erhält man den entsprechenden Radius durch Multiplikation mit der Vergrößerung M :

$$r = M C_s \alpha_0^3 \quad (5.2)$$

Der in Abbildung 5.1 dargestellte Wegunterschied Δ führt zu der oben erwähnten Phasenverschiebung zwischen axialen und achsfernen Strahlen.

Ebenso läßt sich eine vom Einfallswinkel α_0 abhängige Phasenverschiebung zwischen axialen und achsfernen Strahlen erreichen, wenn man die *Anregung* der Linse ändert. Dies bedeutet, daß man die Brechkraft der Linse variiert, indem man den Linsenstrom ändert:

$$f \longrightarrow f + \Delta f \quad \text{mit} \quad \begin{cases} \Delta f < 0 & \text{Überfokussieren} \\ \Delta f > 0 & \text{Unterfokussieren} \end{cases} \quad (5.3)$$

Bei festgehaltener Objekt- und Bildweite werden die von einem Punkt P ausgehenden Strahlen im Falle des Überfokussierens bei Abwesenheit von sphärischer Aberration *vor* der Gaußschen Bildebene zu einem Punkt vereinigt, im Falle des Unterfokussierens *hinter* der Gaußschen Bildebene. Genau derselbe Effekt läßt sich auf eine zweite Art erreichen, indem man bei festgehaltener Bildebene und festgehaltenem Objektivstrom die Gegenstandsweite a variiert. Vergrößern der Gegenstandsweite von $a \longrightarrow a + \Delta a$ entspricht einer Überfokussierung $\Delta f = -\Delta a$.

Die Winkelabweichung ϵ vom idealen Gaußschen Fall verursacht durch sphärische Aberration, Änderung der Gegenstandsweite sowie der Fokussierung ist in Abb. 5.2 jeweils gesondert dargestellt. Die aus diesen drei Teileffekten entstehende Gesamtwinkelabweichung läßt sich durch wenige elementare geometrische Umformungen für den Fall $a \simeq f$ folgendermaßen darstellen [39]:

$$\epsilon = \epsilon_s + \epsilon_a + \epsilon_f = C_s \frac{R^3}{f^4} + (\Delta a - \Delta f) \frac{R}{f^2} \quad (5.4)$$

Hierbei ist R der Abstand zwischen optischer Achse und dem Punkt, an dem ein Strahl die Linse durchquert. Der optische Wegunterschied ds zu einem Strahl im idealen Gaußschen Fall ergibt sich (Abb. 5.3) zu:

$$ds = \epsilon dR \quad (5.5)$$

Durch Integration von (5.4) über den Linsenradius erhält man dann für die Phasenverschiebung $\chi(\theta)$:

$$\chi(\theta) = \frac{2\pi}{\lambda} \Delta = \frac{2\pi}{\lambda} \int_0^R \epsilon dR = \frac{2\pi}{\lambda} \left[\frac{1}{4} C_s \frac{R^4}{f^4} + \frac{1}{2} (\Delta a - \Delta f) \frac{R^2}{f^2} \right] \quad (5.6)$$

Indem man $R/f \simeq \theta$ (siehe Abb. 5.2) setzt und im sogenannten *Defokus* $\Delta z = \Delta a - \Delta f$ den Effekt der Änderung von Brennweite und Gegenstandsweite zusammenfaßt, ergibt sich für die Phasenverschiebung $\chi(\theta)$

$$\chi(\theta) = \frac{\pi}{2\lambda} (C_s \theta^4 + 2\Delta z \theta^2) \quad (5.7)$$

Durch die Einführung der den einzelnen Strahlen zugeordneten Raumfrequenzen $k = \theta/\lambda$ wird daraus:

$$\chi(k) = \frac{\pi}{2} (C_s \lambda^3 k^4 + 2\Delta z \lambda k^2) \quad (5.8)$$

Dieser Ausdruck für die Phasenverschiebung $\chi(k)$ in Gleichung 5.8 wird Aberrationsfunktion genannt. Durch Einführung generalisierter Größen läßt sich die Aberrationsfunktion unabhängig von der sphärischen Aberration und von der Wellenlänge darstellen. Mit den dimensionslosen Größen θ^* , bzw. k^* und Δz^*

$$\theta^* = \theta (C_s/\lambda)^{1/4} = k^* = k (C_s \lambda^3)^{1/4} \quad \text{und} \quad \Delta z^* = \Delta z (C_s \lambda)^{-1/2} \quad (5.9)$$

erhält man für die Aberrationsfunktion den Ausdruck:

$$\chi(\theta^*) = 2\pi \left[\frac{\theta^{*4}}{4} + \frac{\theta^{*2}}{2} \Delta z^* \right] \quad (5.10)$$

Eine Darstellung dieser Funktion findet sich in Abb. 5.4. Ohne Rückübersetzung auf die gerätespezifischen Parameter C_s und λ sei vorausgeschickt, daß die dort dargestellten Phasenverschiebungen in der Größenordnung von mehreren hundert Grad – je nach Beugungswinkel und Defokus – typisch für die hochauflösende Abbildung sind. Dies bedeutet, daß die Austrittswellenfunktion auf höchst drastische Art und Weise, die intuitiv nicht mehr nachvollziehbar ist, modifiziert wird. Dies sei an einem kleinen Beispiel für ein schwaches Phasenobjekt demonstriert:

Die Austrittswellenfunktion setze sich aus dem Nullstrahl und zwei zum Koordinatenursprung symmetrischen abgelenkten Wellen zusammen. Des weiteren wird eine Phasenverschiebung von $\pi/2$ der abgelenkten Strahlen gegenüber dem Nullstrahl angenommen:

$$\Psi(\vec{r}) = A_0 \exp(2\pi i \vec{k}_0 \vec{r}) + A_g \exp(i\pi/2) \left[\exp(2\pi i (\vec{k}_0 + \vec{g}) \vec{r}) + \exp(2\pi i (\vec{k}_0 - \vec{g}) \vec{r}) \right] \quad (5.11)$$

Zuerst wird die Bildintensität $I(\vec{r})$ ohne Einbeziehung der Aberrationsfunktion betrachtet. Sie ergibt sich durch Betragsquadratbildung $\Psi\Psi^*$:

$$I(\vec{r}) = A_0^2 + \underbrace{4A_g^2 \cos^2 2\pi \vec{g} \vec{r}}_{\simeq 0} + 4A_0 A_g \underbrace{\cos \frac{\pi}{2}}_{=0} \cos 2\pi \vec{g} \vec{r} \quad (5.12)$$

Obige Gleichung sagt aus, daß fast gar kein Kontrast entsteht:

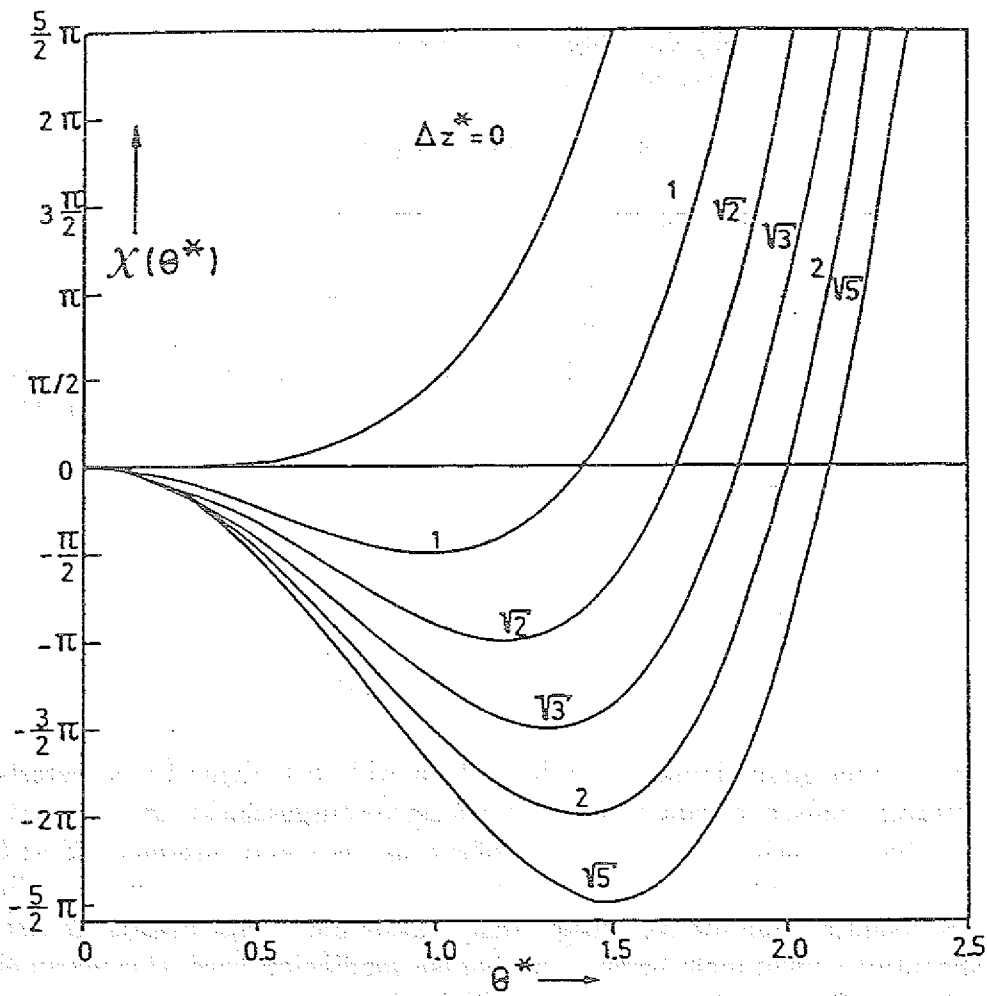


Abbildung 5.4: Darstellung der Aberrationsfunktion $\chi(\theta^*)$ als Funktion des generalisierten Beugungswinkels θ^* für sechs verschiedene Defokuswerte Δz . Bis auf $\Delta z = 0$ alle Unterfokus.

- Der erste Term auf der rechten Seite, der auf den Nullstrahl zurückzuführen ist, hängt nicht vom Ort ab. Er erzeugt im Bild gleichmäßige Helligkeit der Intensität A_0^2 , jedoch keinen Kontrast.
- Der zweite Term ist wegen der Annahme eines *schwachen* Phasenobjekts ($A_g^2 \ll 1$) nahezu null und trägt kaum zum Kontrast bei. Überdies ist dieser Term für eine *Bildfrequenz* verantwortlich, die in der Austrittswellenfunktion nicht vorkommt. Dieser Term erzeugt – abgesehen von seiner geringen Intensität – im Bild Artefakte. Er läßt sich folgendermaßen umformen:

$$\cos^2 2\pi \vec{g} \vec{r} = 1/2 (1 + \cos 4\pi \vec{g} \vec{r}) \quad (5.13)$$

Im Gegensatz zur Austrittswellenfunktion, die nur einen Term $\sim \cos 2\pi \vec{g} \vec{r}$ beinhaltet, tritt hier die doppelte Bildfrequenz auf.

- Der dritte Term ist zwar $\sim \cos 2\pi \vec{g} \vec{r}$ und beinhaltet die Raumfrequenz g , die die Präparatstruktur charakterisiert. Wegen des Phasensprungs der abgebeugten Strahlen um $\pi/2$ ist er jedoch exakt null.

Um überhaupt Bildkontrast zu erhalten, ist die zusätzliche Phasenverschiebung, die die Aberrationsfunktion bewirkt, also unbedingt nötig. Mit Einbeziehung der Aberrationsfunktion als zusätzliche Phasenverschiebung

$$\frac{\pi}{2} \longrightarrow \frac{\pi}{2} + \chi(\vec{g}) \quad (5.14)$$

erhält man für den dritten Term in Gleichung (5.12):

$$I_g(\vec{r}) = 4A_0A_g \cos 2\pi \vec{g} \vec{r} \cos(\pi/2 + \chi(\vec{g})) = 4A_0A_g \cos 2\pi \vec{g} \vec{r} \underbrace{\sin \chi(\vec{g})}_{\equiv B_C} \quad (5.15)$$

Die sogenannte *kohärente Kontrasttransferfunktion* $B_C = \sin \chi$ bestimmt also, in welchem Maße die Raumfrequenz g vom Mikroskop übertragen wird. Der Kontrast ist gerade dann maximal, wenn $B_C = \pm 1$, d.h. wenn

$$\chi(\vec{g}) = (2n - 1) \frac{\pi}{2} \quad (5.16)$$

Die kontrastbildende Wirkung durch eine Phasendrehung um $\pm\pi/2$ ist in Abb. 5.5 noch einmal anschaulich verdeutlicht. Ebenso treten Kontrastmaxima bei Phasendrehungen, deren Betrag ein ungeradzahliges Vielfaches von $\pi/2$ ist, auf. Reduzierte Defokuswerte, wo die Aberrationsfunktion gerade bei diesen Werten ein Minimum hat, sind in Abb. 5.4 dargestellt. Sie ergeben sich zu $\Delta z^* = -\sqrt{2n-1}$ mit $n \geq 1$. In der Nähe der Minima, wo die Steigung der Aberrationsfunktion am geringsten ist, werden benachbarte Raumfrequenzbereiche mit etwa derselben Phasenverschiebung versehen. Der Bildkontrast ist über einen weiten Raumfrequenzbereich also gleich „gut“. Man spricht in diesem Falle von sogenannten *Paßbändern*. Der Defokuswert $\Delta z^* = -1$ hat den Vorteil, daß das Minimum dort besonders breit ist, also ein großer Frequenzbereich mit etwa derselben Phasendrehung zur

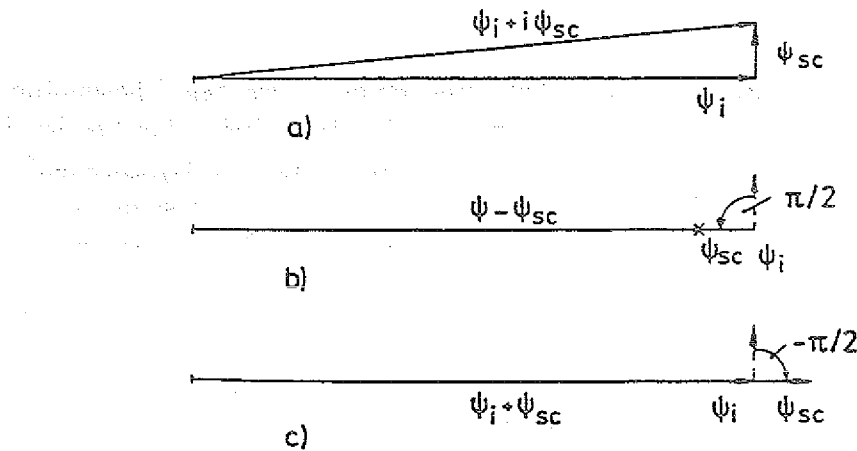


Abbildung 5.5: Darstellung des Einflusses der zusätzlichen Phasendrehung einer abgebeugten Welle Ψ_{sc} relativ zur einfallenden Welle Ψ um $\pm\pi/2$ in der komplexen Zahlenebene. Im Fall (a) unterscheidet sich der Betrag der resultierenden Welle kaum von dem der einfallenden Welle. (b) Positiver Phasenkontrast durch Drehung um $\pi/2$, (c) negativer Phasenkontrast.

Kontrastbildung beiträgt. Noch deutlicher wird dieses Verhalten bei Betrachten der Kontrasttransferfunktion.

Eine Darstellung der kohärenten Kontrasttransferfunktion (CTF) des JEOL 4000EX für verschiedene Defokuswerte befindet sich in Abb. 5.6. Deutlich zu sehen sind die oben beschriebenen Paßbänder. Die schnelle Oszillation der Kurven zu großen Raumfrequenzen hin wird durch die starke Steigung der Aberrationsfunktion (Abb. 5.4) aufgrund des Terms $C_s k^4$ bewirkt. In der Nähe von $\Delta z^* = -1$ befindet sich der sogenannte *Scherzerfokus*, bei dem die Phasenverschiebung über einen weiten Frequenzbereich etwa $2\pi/3$ beträgt (Abb. 5.6 c). Durch die Abweichung vom Idealwert $\pi/2$ erreicht man eine größere Ausdehnung des Bereichs, an dem die CTF nahezu 1 ist, zu großen Raumfrequenzen hin. Damit wird im Falle des *schwachen Phasenobjekts* eine kontinuierliche Kontrastübertragung bis zu kleinen Objektdetails gewährleistet.

Setzt man als Abschätzung den ungefähren Wert von $2\pi/3$ (eine genaue Herleitung findet sich in [49]) in Gleichung (5.8) ein, erhält man für den Defokus:

$$\Delta z = -(4C_s \lambda / 3)^{1/2} \quad (5.17)$$

Der Nulldurchgang der CTF liegt dann bei der Raumfrequenz

$$g_0 = (3C_s \lambda^3 / 4)^{-1/4} \quad (5.18)$$

Für größere Raumfrequenzen erhält man schnelle Oszillationen der CTF, die keine direkte Bildinterpretation mehr zulassen. Schließt man diese durch eine Objektiv-

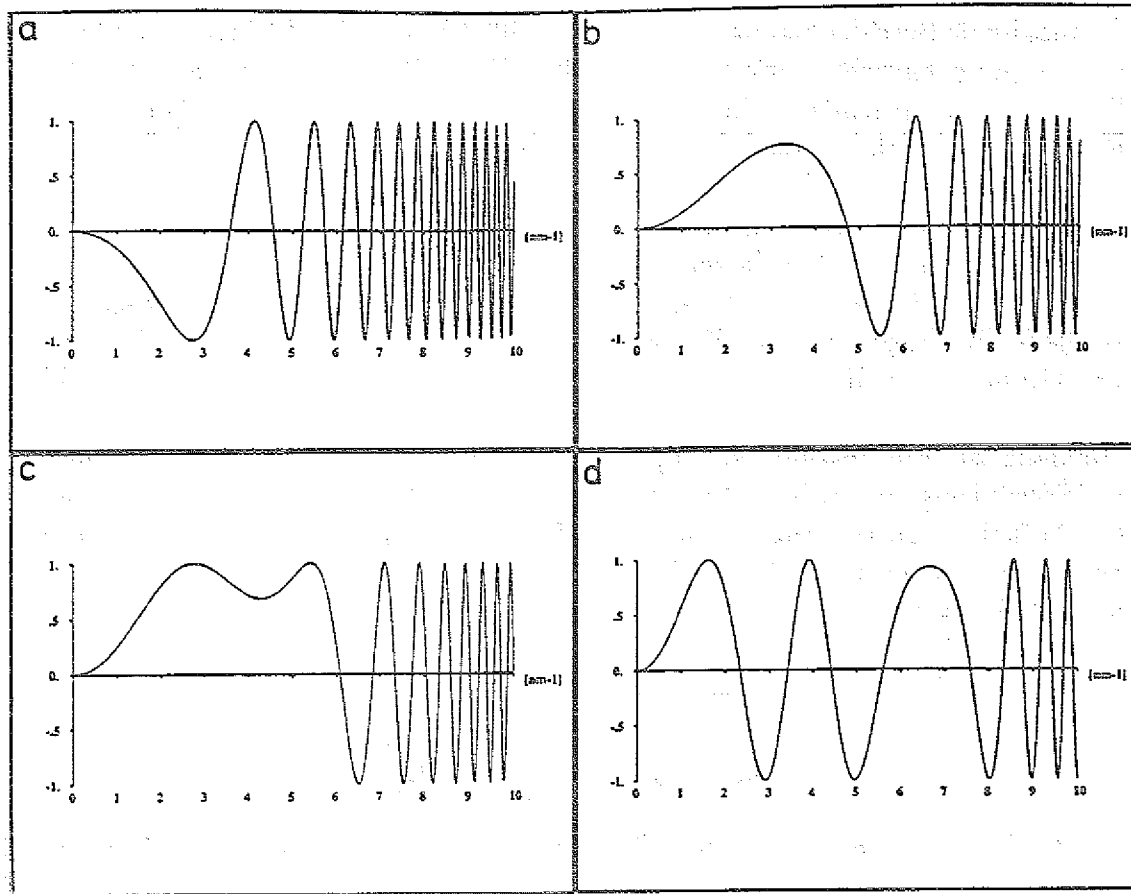


Abbildung 5.6: Darstellung der kohärenten Kontrasttransferfunktion $B_C(k)$ des JEOL 4000EX für die Defokuswerte: a) 30, b) -30, c) -50, d) -120 nm; Parameter: $C_s = 1.0$ mm, $E = 400$ kV.

blende aus, so ergibt sich für den kleinsten direkt interpretierbaren Abstand:

$$d = 1/g_0 = (3/4)^{1/4} (C_s \lambda^3)^{1/4} \quad (5.19)$$

Die genaue Darstellung ergibt für die sogenannte *Punktauflösung* des Instruments:

$$d_{\min} = 0.66 (C_s \lambda^3)^{1/4} \quad (5.20)$$

Das Punktauflösungsvermögen des JEOL 4000EX beträgt 1.7 nm. Dabei wurde die maximale Beschleunigungsspannung von 400 kV, der eine Elektronenwellenlänge von 1.6 pm entspricht, sowie ein C_s von (laut Hersteller) 1.0 mm vorausgesetzt. Das Punktauflösungsvermögen dient eher als Vergleichsmaßstab für die Leistung eines Elektronenmikroskops, da man es in der Realität selten mit schwachen Phasenobjekten zu tun hat. Zudem wird bei dieser Betrachtung der Bereich jenseits des Nulldurchgangs der CTF mit $g > g_0$, dem noch feinere Objektdetails entsprechen, vollkommen außer Acht gelassen.

Neben der Frage, bei welchem Defokuswert „optimaler“ Kontrast gegeben ist, verdient die Frage nach der Bedingung für minimalen Kontrast noch Beachtung: Der Grund hierfür liegt in der Tatsache, daß der im Experiment benutzte Defokuswert von der Anregung der Objektivlinse *und* vom Abstand Objekt – Objektivlinse abhängig ist. Ein bestimmter Objektivlinsenstrom ist also nicht einer bestimmten Defokussierung zuzuordnen. Eine Möglichkeit, den Defokuswert einer experimentellen Aufnahme zu bestimmen, ist das Auftreten minimalen Bildkontrastes an einer amorphen Probenstelle, die ein Kontinuum an Raumfrequenzen anbietet, als Anhaltspunkt zu benutzen. Die Fragestellung lautet also: Bei welcher Defokussierung tritt minimaler Kontrast auf?

Von Gleichung (5.16) ausgehend erhält man als Bedingung für maximalen Kontrast:

$$\Delta z = (2n - 1) \frac{1}{2\lambda g^2} - \frac{C_s \lambda^2 g^2}{2} \quad (5.21)$$

Hierbei betrachtet man bei festgehaltenem n den Defokus für optimalen Kontrast als Funktion der Raumfrequenz g . Verglichen mit dem Fall einer idealen Linse ($C_s = 0$) sind diese Kurven aufgrund des Terms

$$\Delta z_s = -\frac{C_s \lambda^2 g^2}{2} \quad (5.22)$$

in Richtung Unterfokus verschoben. Dem in [50] vorgeschlagenen Weg zur Auffindung der Bedingung für *minimalen* Kontrast liegt folgender Gedankengang zugrunde: Grundsätzlich erfüllt jedes *ungerade halbzahlige* n in Gleichung (5.21) die Bedingung minimalen Kontrasts. Es gilt nun, denjenigen Raumfrequenzbereich zu ermitteln, über dem die Δz Kurven (5.21) am wenigsten variieren, also ein möglichst breites Band charakterisiert wird, das mit minimalem Kontrast abgebildet wird. Für $n=0$ hat Δz ein Maximum an der Stelle g_0 , die Kurven für $n \geq 1/2$ besitzen jedoch nur einen Wendepunkt. An diesen Stellen variieren die Kurven für Δz am wenigsten. In [50] wird das Maximum bei g_0 für $n = 0$, das im Vergleich zu den Bereichen um die Wendepunkte über einen relativ breiten Raumfrequenzbereich wenig variiert, als

Kriterium herangezogen. ($n = 0$ entspricht einer Phasenverschiebung von $-\pi/2$) Für dieses Kontrastmaximum gilt:

$$\frac{d[\Delta z(n=0)]}{dg} = \frac{1}{\lambda g^3} - C_s \lambda^2 g = 0 \quad (5.23)$$

Damit findet man für g_0 :

$$g_0 = (C_s \lambda^3)^{-1/4} \quad (5.24)$$

Durch Einsetzen von (5.24) in (5.22) erhält man das nächstliegende Kontrastminimum mit $n = 1/2$, dem die Phasenverschiebung 0 entspricht, bei einem Defokus

$$\Delta z_{\min} = -\frac{1}{2} (\lambda C_s)^{1/2} \quad (5.25)$$

Für das JEOL 4000EX ($C_s = 1.0 \text{ mm}$, $\lambda(400 \text{ kV}) = 1.6 \text{ pm}$) ergibt sich das Kontrastminimum nach (5.25) bei einem Unterfokus von -20.3 nm .

Eine weitere wichtige Eigenschaft der kohärenten CTF besteht darin, daß der *lineare* Kontrastbeitrag einer bestimmten Raumfrequenz g sich mit dem Defokus Δz periodisch wiederholt. Aus der Tatsache, daß der Wert der Aberrationsfunktion χ nur modulo 2π in die CTF eingeht, erhält man mit

$$\chi(\vec{g}, \Delta z_2) - \chi(\vec{g}, \Delta z_1) = 2\pi \quad (5.26)$$

aus Gl. (5.8) für jede Raumfrequenz g eine *unterschiedliche* Wiederholungsperiode $\Delta z_2 - \Delta z_1$:

$$\Delta z_2 - \Delta z_1 = 2/\lambda g^2 \quad (5.27)$$

5.1.2 Kontrast bei partiell kohärenter Einstrahlung

In den bisherigen Ausführungen wurde der einfallende Elektronenstrahl als ideale ebene Welle behandelt. Dieses Modell weicht jedoch erheblich von den realen Gegebenheiten in einem Elektronenmikroskop ab und bedarf einiger einschneidender Korrekturen, um die experimentellen Ergebnisse zu beschreiben. Die im vorangehenden Abschnitt beschriebene kohärente CTF erstreckt sich zu beliebig großen Raumfrequenzen. Dies würde einer Übertragung von Informationen bis hinunter zur Wellenlänge des einfallenden Elektronenstrahls bedeuten, die lediglich durch die schnellen Oszillationen der CTF keine direkte Bildinterpretation mehr erlauben. Grundsätzlich müßte diese Information jedoch mittels numerischer Bildsimulation zugänglich sein. Aufgrund des Einflusses von nur *partieller räumlicher* und *partieller zeitlicher Kohärenz* des einfallenden Strahls ist das *Informationslimit* eines Elektronenmikroskops um etwa 3 Größenordnungen schlechter als die entsprechende Wellenlänge. Im folgenden wird die Auswirkung eines räumlich nur partiell kohärenten Strahls auf das Auflösungsvermögen eines Mikroskops ausführlich dargestellt. Mit dem dabei gewonnen Formalismus läßt sich die zeitliche Inkohärenz auf einfache Art analog behandeln. Überdies werden alle wichtigen Elemente der Bildentstehung für Objekte, die nicht als schwache Phasenobjekte beschrieben werden können – was dem allgemeinen Fall entspricht – eingeführt.

Partielle räumliche Kohärenz

Mit dem Begriff *partielle räumliche Kohärenz* trägt man der Tatsache Rechnung, daß der Elektronenstrahl in der Nähe der Objektebene fokussiert werden muß, um ausreichende Intensität zu erhalten. Man hat also in der Realität einen konvergenten Elektronenstrahl, dessen Wellenfronten im Gegensatz zu einer ebenen Welle gekrümmt sind. Hier ist ein einfaches Gedankenmodell sehr dienlich, um sich die Auswirkungen dieser Strahlform klarzumachen: Man kann sich den einfallenden Strahl aus „Substrahlen“ zusammengesetzt denken, die einen Kegel ausfüllen, dessen Achse parallel zur optischen Achse des Mikroskops ausgerichtet ist. Die Spitze des Kegels liege in der Nähe der Probenoberfläche. Diese Substrahlen denke man sich von einer *effektiven Quelle* [51, 52, 53] ausgehend, die der Grundfläche dieses Kegels entspricht. Die Grundfläche des Kegels und somit der Emissionsort dieser „Substrahlen“ läßt sich mit der Kondensorblendenöffnung identifizieren. Diese Strahlen kommen je nach ihrer Einfallsrichtung aufgrund der unterschiedlichen optischen Weglängen in der Objektebene mit unterschiedlichen Phasen an. Man hat also ein Kontinuum an einfallenden Strahlen, die jeder für sich ein Beugungsbild erzeugen. Die von diesen „Substrahlen“ erzeugten Beugungsbilder sind natürlich um den jeweiligen Strahl zentriert. Dies entspricht einer Translation dieser Beugungsbilder in der hinteren Brennebene der Objektivlinse. Im Falle einer zur optischen Achse zentrosymmetrischen Quelle (Kegel) sind diese Translationen im Fourierraum ebenfalls zentrosymmetrisch um die Lage der δ -förmigen Intensitätsmaxima, die eine ebene Welle erzeugen würde. Auf diese Weise wird aus einem Beugungspunkt eine Beugungsscheibe. Je stärker der Elektronenstrahl mittels des Kondensorsystems konvergiert wird, umso breiter werden die Beugungsscheiben im Fourierraum. Die Verbreiterung der Beugungspunkte im Fourierraum ist also eine Folge der Benutzung eines Kontinuums von einfallenden Strahlen bzw. ebenen Wellen, deren Einfallsrichtungen symmetrisch um die optische Achse verteilt sind.

Diese Verteilung bzw. Gewichtung der einzelnen ebenen Wellen aus dem Kontinuum kann durch eine um die optische Achse radialsymmetrische Funktion dargestellt werden. Da diese Teilwellen aufgrund ihres Wegunterschiedes untereinander *nicht* mehr interferenzfähig sind, müssen deren Beiträge zum elektronenmikroskopischen Bild – mit der oben beschriebenen Verteilungsfunktion gewichtet – inkohärent aufintegriert werden. Man kann sich vorstellen, daß jede Teilwelle für sich aufgrund von Interferenz ein eigenes Bild erzeugt, dessen Intensität mit den Bildern aller anderen Teilwellen *ohne* Interferenzeffekte überlagert wird. Der Gewichtungsfaktor sagt aus, wie stark eines der Teilbilder zum Gesamtbild beiträgt. Man erkennt intuitiv, daß eine Überlagerung vieler ähnlicher Bilder zu einer „Verschmierung“ desjenigen Bildes führt, das man bei idealer räumlich kohärenter Einstrahlung erhalten würde. Diese „Verschmierung“, die zu einer Begrenzung des Auflösungsvermögens führt, soll im folgenden mathematisch beschrieben werden.

Die oben erwähnte Intensitätsverteilungsfunktion hängt vom Winkel θ , den die einzelnen einfallenden Teilstrahlen mit der optischen Achse bilden, ab. Diesem Winkel kann man über die Gleichung $|\vec{q}| = \theta/\lambda$ einen reziproken Vektor $\vec{q}$ zuordnen. Der Vektor $\vec{q}$ liegt *innerhalb* einer Beugungsscheibe. Er definiert den vom Mittelpunkt

der Beugungsscheibe aus gemessenen Ort des infinitesimalen Intensitätsbeitrages einer bestimmten Teilwelle.

Die mathematische Behandlung der partiellen räumlichen Kohärenz der Elektronenquelle soll nun erläutert werden. Die Austrittswellenfunktion werde beschrieben durch

$$\Psi(\vec{r}) = \int F(\vec{k}) \exp(-2\pi i \vec{k} \vec{r}) d\vec{k}, \quad (5.28)$$

wobei mit $F(\vec{k})$ die komplexen Fourierkoeffizienten der Wellenfunktion $\Psi(\vec{r})$ bezeichnet werden:

$$F(\vec{k}) = \int \Psi(\vec{r}) \exp(2\pi i \vec{k} \vec{r}) d\vec{r} \quad (5.29)$$

Im Falle eines idealen Mikroskops erhält man die Bildintensität $I(\vec{k})$ über die Autokorrelation der Wellenfunktion im Fourierraum (Siehe Anhang A):

$$I(\vec{k}) = \int F(\vec{k}' + \vec{k}) F^*(\vec{k}') d\vec{k}' \quad (5.30)$$

Die Berücksichtigung eines realen Mikroskops umfaßt nun zwei Effekte:

1. Die in der obigen Gleichung auftretenden Koeffizienten F und F^* müssen mit den durch die Objektivlinse hervorgerufenen Phasenverschiebungen $\chi(\vec{k})$ beaufschlagt werden. Dies geschieht durch einfaches Heranmultiplizieren des Exponentialterms $P(\vec{k}) = \exp(-2\pi i \chi(\vec{k}))$ an die Fourierkoeffizienten. Dieser Vorgang beschreibt die im vorangehenden Abschnitt besprochene *kohärente* Bildentstehung bei Einstrahlung einer ebenen Welle.
2. Darüberhinaus soll die einfallende ebene Welle gemäß dem Modell der „Substrahlen“ kontinuierlich um die optische Achse gekippt werden. Die von den jeweiligen „Substrahlen“ erzeugten Bilder sollen dann *inkohärent* zur Gesamtbildintensität aufintegriert werden.

Zu diesem Zwecke wird – um die Anschauung zu erhöhen – rein formal *vor* der Integration in (5.30) die für eine Autokorrelation typische Beziehung der Argumente des Integranden außer Acht gelassen. Solange die Integration nicht wirklich durchgeführt wird, ist diese Manipulation unerheblich. Die Argumente aus (5.30) werden folgendermaßen substituiert:

$$F(\vec{k}' + \vec{k}) F^*(\vec{k}') \longrightarrow F(\vec{\alpha}) F^*(\vec{\beta}) \quad (5.31)$$

Nun wird der oben angeführte Punkt 1, die Berücksichtigung der kohärenten CTF mittels des Terms $P = \exp(-2\pi i \chi)$ vollzogen. Damit wird aus dem Argument von (5.30):

$$F(\vec{\alpha}) P(\vec{\alpha}) F^*(\vec{\beta}) P^*(\vec{\beta}) \quad (5.32)$$

Punkt 2 umfaßt die kontinuierliche Strahlkipfung, die – wie oben beschrieben – durch einen Vektor $\vec{q}$, der innerhalb einer Beugungsscheibe liegt, dargestellt wird. Die daraus resultierenden Intensitätsbeiträge werden mit der Verteilungsfunktion $s(\vec{q})$ gewichtet und über den Flächeninhalt der Beugungsscheiben integriert. Die

Verschiebung des Beugungsbildes um den Vektor $\vec{q}$ bei Kippung des Strahls um den Winkel $\theta = |\vec{q}|\lambda$ ist an dem nun neuerlich modifizierten Ausdruck (5.33) einfach zu sehen:

$$\int s(\vec{q}) F(\vec{\alpha} + \vec{q}) P(\vec{\alpha} + \vec{q}) F^*(\vec{\beta} + \vec{q}) P^*(\vec{\beta} + \vec{q}) d\vec{q} \quad (5.33)$$

Das Integral beschreibt die *inkohärente* Überlagerung der zu den „Substrahlen“ gehörenden Teilbilder. Die Verteilungsfunktion $s(\vec{q})$ beschreibt die Gewichtung der einzelnen Bilder. Dieser Ausdruck ist analytisch exakt, jedoch in der Realität auch mit unverhältnismäßig großem Rechenaufwand numerisch nur näherungsweise zu bewältigen. Der Grund hierfür ist, daß man die Fourierkoeffizienten F der Austrittswellenfunktion für ein Kontinuum von Argumenten $\vec{q}$, d.h. für ein Kontinuum von einfallenden Strahlen berechnen müßte. In den heute gängigen Simulationsprogrammen für hochauflösende Elektronenmikroskopie wird der Effekt eines konvergenten Strahls auf die dynamische Beugung vernachlässigt. Der Strahl wird für den Durchgang durch das Objekt als parallel angenommen, der Konvergenz wird erst bei den elektronenoptischen Vorgängen in der Objektlinse Rechnung getragen. Dies äußert sich mathematisch im Vorziehen der Fourierkoeffizienten vor das Integral in (5.33):

$$F(\vec{\alpha}) F^*(\vec{\beta}) \int s(\vec{q}) P(\vec{\alpha} + \vec{q}) P^*(\vec{\beta} + \vec{q}) d\vec{q} \quad (5.34)$$

Zuletzt muß noch berücksichtigt werden, ob eine Beugungsscheibe innerhalb der Objektivblende, teilweise innerhalb oder gänzlich außerhalb liegt. Dies geschieht mittels der Aperturfunktion $a(\vec{k})$:

$$a(\vec{k}) = \begin{cases} 1 & \text{wenn } |\vec{k}| \leq k_0 \\ 0 & \text{sonst} \end{cases} \quad (5.35)$$

Hierbei ist k_0 der Blendenradius im Fourierraum. Durch Erweiterung der Funktion $P = \exp(-2\pi i \chi)$ (5.34) zur Funktion $t(\vec{k})$ mit

$$t(\vec{k}) = \exp(-2\pi i \chi) a(\vec{k}) \quad (5.36)$$

wird aus (5.34)

$$F(\vec{\alpha}) F^*(\vec{\beta}) \underbrace{\int s(\vec{q}) t(\vec{\alpha} + \vec{q}) t^*(\vec{\beta} + \vec{q}) d\vec{q}}_{T(\alpha, \beta)} \quad (5.37)$$

Ersetzt man gemäß (5.31) die Argumente $\vec{\alpha}, \vec{\beta}$ durch die ursprünglichen, so erhält man für die Bildintensität $I(\vec{k})$ unter Berücksichtigung partieller räumlicher Kohärenz:

$$I(\vec{k}) = \int T(\vec{k}' + \vec{k}, \vec{k}') F(\vec{k}' + \vec{k}) F^*(\vec{k}') d\vec{k}' \quad \text{mit} \quad (5.38)$$

$$T(\vec{k}', \vec{k}'') = \int s(\vec{q}) t(\vec{k}' + \vec{q}) t^*(\vec{k}'' + \vec{q}) d\vec{q} \quad (5.39)$$

Die Funktion $T(\vec{k}', \vec{k}'')$ wird in der englischsprachigen Literatur mit *transmission cross coefficient* bezeichnet. In dieser Arbeit soll sie schlicht als Kreuzkoeffizient bezeichnet werden. Wie an späterer Stelle deutlicher zu sehen sein wird, legt der Kreuzkoeffizient sowohl die Phasendrehungen zweier Strahlen $\vec{k}', \vec{k}''$ als auch die Dämpfung ihres gemeinsamen Interferenzbeitrages zur Gesamtbildintensität fest.

Mit einer Aufteilung der Wellenfunktion in Nullstrahl und abgebeugte Strahlen

$$F(\vec{k}) = \delta(\vec{k}) + O(\vec{k}) \quad (5.40)$$

erhält man durch Einsetzen in (5.38) für die Bildintensität $I(\vec{k})$:

$$\begin{aligned} I(\vec{k}) &= \int T(\vec{k}' + \vec{k}, \vec{k}') \{ \delta(\vec{k}' + \vec{k}) + O(\vec{k}' + \vec{k}) \} \{ \delta(-\vec{k}') + O^*(\vec{k}') \} d\vec{k}' \\ &= T(\vec{k}, 0) \delta(\vec{k}) \\ &\quad + \{ T(0, -\vec{k}) O^*(-\vec{k}) + T(\vec{k}, 0) O(\vec{k}) \} \\ &\quad + \int T(\vec{k}' + \vec{k}, \vec{k}') O(\vec{k}' + \vec{k}) O^*(\vec{k}') d\vec{k}' \end{aligned} \quad (5.41)$$

Die Bildintensität $I(\vec{k})$ läßt sich gemäß (5.41) in 3 fundamentale Anteile aufspalten:

- Der erste Term stellt den Anteil des Nullstrahls dar und erzeugt ortsunabhängigen Hintergrund der Intensität 1.
- Der zweite Term repräsentiert alle Intensitätsanteile im Bild, die auf Interferenz zwischen Nullstrahl und einem der abgebeugten Strahlen zurückzuführen sind.
- Der dritte Term umfaßt alle Intensitätsanteile, die durch Interferenz der abgebeugten Strahlen untereinander zustandekommen.

Der erste Term stellt die Bildintensität dar, die die Elektronenquelle ohne abzubildendes Objekt auf dem Leuchtschirm erzeugen würde. Das Einfügen eines „transparenten“ Objekts äußert sich in der von 1 abweichenden Intensitätsmodulation durch die beiden Interferenzterme in (5.41).

Geht man vom Idealfall des *schwachen* Phasenobjekts aus, spielt der letzte Interferenzterm bei der Bildentstehung keine Rolle, da er ausschließlich die Fourierkoeffizienten $O(\vec{k})$ in zweiter Ordnung enthält. Diese Terme zweiter Ordnung sind wegen der Annahme geringer Intensität abgebeugter Strahlen im Verhältnis zu Termen, die den Nullstrahl enthalten, zu vernachlässigen. Die Abbildung eines schwachen Phasenobjekts ist aus diesem Grunde ausreichend durch den Intensitätsbeitrag

$$I_2(\vec{k}) = T(0, -\vec{k}) O^*(-\vec{k}) + T(\vec{k}, 0) O(\vec{k}) \quad (5.42)$$

beschrieben. Die ausschließliche Berücksichtigung dieses Terms wird auch *lineare* Bildberechnung genannt, da das Objektspektrum nur in erster Ordnung auftritt.

Um die Bildintensität nach (5.42) berechnen zu können, fehlt noch die Information über das genaue Aussehen der Intensitätsverteilungsfunktion $s(\vec{q})$ der effektiven Quelle, die in die Berechnung des Kreuzkoeffizienten eingeht. Wie eingangs schon beschrieben, ist zumindest im gewünschten Idealfall die Intensität der Quelle symmetrisch zur optischen Achse verteilt. Damit ist s nur mehr von $|\vec{q}|$ abhängig, bzw. durch einen Winkel θ zur optischen Achse darzustellen. Bleibt man bei dem eingangs beschriebenen Modell der kegelförmig um die optische Achse verteilten „Substrahlen“ gleicher Intensität, so bedeutet dies konstante Intensität bis zum Grenzwinkel θ_0 , der dem Öffnungswinkel des Kegels entspricht. Für Winkel größer θ_0 sei die abgestrahlte Intensität exakt null. Mit der Forderung, daß die Gesamtintensität der Quelle zu 1 normiert sein soll, erhält man nach Umsetzung der Winkel in reziproke Vektoren für die Intensitätsverteilung der effektiven Quelle:

$$s(|\vec{q}|) = \begin{cases} (\pi q_0^2)^{-1} & \text{wenn } |\vec{q}| \leq q_0 \\ 0 & \text{sonst} \end{cases} \quad (5.43)$$

Diese Verteilung stellt im Fourierraum einen Zylinder der Höhe $1/\pi q_0^2$ dar, dessen Achse parallel zur optischen Achse ist. Eine zweite einleuchtende Möglichkeit zur Beschreibung der Intensitätsverteilung besteht darin, die Kanten des Zylinders zu glätten, indem man die Zylinderfunktion durch eine Gaußverteilung ersetzt:

$$s(|\vec{q}|) = \frac{1}{\pi q_0^2} \exp\left(-\frac{|\vec{q}|^2}{q_0^2}\right) \quad (5.44)$$

Die Größe $\theta_c = \lambda q_0$, die im Zylindermodell exakt den Öffnungswinkel des Strahls, im Gauß'schen Modell die halbe $1/e$ -Breite der Intensitätsverteilung darstellt, wird Semikonvergenzwinkel genannt. Er beeinflußt nachhaltig die Abbildungsqualität.

Diskussionen zum Konzept der effektiven Quelle sowie Vergleiche der beiden oben dargestellten Modelle zur Intensitätsverteilung finden sich in [52, 51]. Während in diesen Veröffentlichungen beide Modelle als ausreichend genau bezeichnet werden, findet sich in [54] die Vermutung, das Gaußsche Modell eigne sich besser für einen defokussierten Strahl, das Zylindermodell jedoch mehr für die Beschreibung des Abbildungszustands, bei dem das Filament in der Probenebene abgebildet wird. Wohl nicht letztlich wegen seiner einfacheren mathematischen Handhabbarkeit wird oft (z.B. [55], wonach das EMS Programm IM1 (Abb. 4.4) und die für diese Arbeit konzipierten Programme geschrieben sind) dem Gaußschen Modell der Vorzug gegeben.

Um ein tieferes physikalisches Verständnis des Zusammenhangs zwischen Semikonvergenz und Bildqualität zu erhalten, ist es angebracht, die Fouriertransformierte der Quellverteilung $s(\vec{q})$ näher zu untersuchen:

$$\Gamma(\vec{r}_1 - \vec{r}_2) = \int s(\vec{q}) \exp(-2\pi i \vec{q}(\vec{r}_1 - \vec{r}_2)) d\vec{q} \quad (5.45)$$

Die Normierungsbedingung für Γ und s lautet:

$$\int s(\vec{q}) d\vec{q} = \Gamma(0) = 1 \quad (\text{Einheitsintensität}) \quad (5.46)$$

Die sogenannte *Phasenkorrelationsfunktion* $\Gamma(\vec{r}_1 - \vec{r}_2)$ gibt an, in welchem Maße die von einer Quelle abgestrahlten Wellenfronten an den Punkten $\vec{r}_1$ und $\vec{r}_2$ untereinander noch in Phase sind. Der Abstand $|\vec{r}_1 - \vec{r}_2|$, bei dem die Phasenkorrelationsfunktion, deren Maximalwert 1 ist, unter eine willkürliche Schwelle (meist 88 %) gefallen ist, wird als *Kohärenzlänge* Λ bezeichnet. Punkte $\vec{r}_1, \vec{r}_2$, deren Abstand geringer als die Kohärenzlänge ist, werden als kohärent beleuchtet angesehen, solche mit größerem Abstand als inkohärent beleuchtet.

Der Zusammenhang zwischen Phasenkorrelationsfunktion Γ , Quellverteilung s , Kohärenzlänge Λ und der Verteilungsbreite q_0 ist in der folgenden Tabelle für die Extremfälle totaler Kohärenz bzw. totaler Inkohärenz sowie für den Fall partieller Kohärenz dargestellt. (Für den letzten Fall sind die schon erwähnten Modelle der Gaußfunktion bzw. der Zylinderfunktion als Beispiel eingetragen).

Quelle	$\Gamma(\vec{r} - \vec{r}_1)$	$s(\vec{q}) = \mathcal{F}(\Gamma)$	Λ	q_0
Kohärent (ebene Welle)	1	$\delta(\vec{q} - \vec{q}_1)$	∞	0
Partiell kohärent	Bsp.: Gaußfunktion	Gaußfunktion	$0 < \Lambda < \infty$	$0 < q_0 < \infty$
	Besselfunktion	Zylinderfunktion	$0 < \Lambda < \infty$	$0 < q_0 < \infty$
Inkohärent	$\delta(\vec{r} - \vec{r}_1)$	konstant	0	∞

Man kann den Semikonvergenzwinkel ($\theta_c = \lambda q_0$ für kleine q_0) als Breite der Quellverteilung im Fourierraum ansehen, die Kohärenzlänge als Breite der Phasenkorrelationsfunktion im Realraum. Die beiden Verteilungen zeigen *entgegengesetztes* Verhalten bezüglich ihrer Breiten: Je breiter also langreichweitiger die Phasenkorrelationsfunktion ist, desto schärfer ist die Quellverteilung im Fourierraum. Im Falle einer ebenen Welle $\exp(2\pi i \vec{q}_1 \vec{r})$, deren Phasen auch noch im Unendlichen korreliert sind, weist die Quellverteilung nur mehr einen Beitrag beim Wellenvektor $\vec{q}_1$ auf. Im Falle einer partiell kohärenten Quelle mit endlicher Korrelationslänge „weicht“ die Intensitätsverteilung je nach Modell zu einer um den Wellenvektor $\vec{q}_1$ zentrierten Gauß- bzw. Zylinderfunktion auf. Bei einer komplett inkohärenten Quelle, bei der die Phasenkorrelationsfunktion im Realraum sich auf einem beliebigen Punkt $\vec{r}_1$ zusammenzieht, verschwindet jegliche Zentrierung um einen Wellenvektor $\vec{q}_1$ im Fourierraum, die Quellverteilung ist dort isotrop.

Für die Berechnung des linearen Beitrags zur Bildintensität (siehe Gleichung (5.42))

$$I_2(\vec{k}) = T(0, -\vec{k})O^*(-\vec{k}) + T(\vec{k}, 0)O(\vec{k}) \quad (5.47)$$

werden die Fourierkomponenten der Wellenfunktion in Real- und Imaginärteil aufgespalten:

$$O(\vec{k}) = O_1(\vec{k}) + i O_2(\vec{k}) \quad (5.48)$$

Der Realteil O_1 beschreibt den Anteil der Wellenfunktion, der keine Phasenverschiebung zum Nullstrahl erfahren hat, der Imaginärteil O_2 repräsentiert die um 90° dazu gedrehte Komponente. Ist der Imaginärteil der *gesamten* Wellenfunktion 0, so

spricht man von einem *Amplitudenobjekt*, der Bildkontrast kommt nur durch die Amplitudenvariation einzelner Strahlen zustande. Ist hingegen der Realteil 0, entsteht der Bildkontrast durch die 90°-Phasenverschiebung der abgelenkten Strahlen. In diesem Falle hat man es mit einem reinen *Phasenobjekt* zu tun. Mit der Annahme eines reinen Phasenobjektes wird bei symmetrischer Quelle aus (5.47):

$$I_2(\vec{k}) = O_2(\vec{k}) \int s(\vec{q}) a(\vec{q}) a(\vec{q} + \vec{k}) \sin \{2\pi [\chi(\vec{q} + \vec{k}) - \chi(\vec{q})]\} d\vec{q} \quad (5.49)$$

Der Übersichtlichkeit halber wird im folgenden der Aperturterm $a(\vec{q}) a(\vec{q} + \vec{k})$ gleich 1 gesetzt, was einfach einer Abbildung mit „genügend“ großer Blende entspricht. Des weiteren beschränkt sich die Diskussion auf „kleine“ Verteilungsbreiten q_0 , bzw. kleine Semikonvergenzwinkel, die die Bedingung $q_0 \ll |\vec{k}|$ erfüllen. Damit läßt sich die Aberrationsfunktion unter Vernachlässigung höherer Terme als der ersten Ordnung in $\vec{q}$ um den Vektor $\vec{k}$ entwickeln:

$$\chi(\vec{k} + \vec{q}) \approx \chi(\vec{k}) + \vec{q} \nabla_{\vec{k}} \chi(\vec{k}) + \dots \quad (5.50)$$

Vernachlässigt man in (5.49) ebenso $\chi(\vec{q})$, da es nur aus Termen zweiter und vierter Ordnung in $\vec{q}$ besteht, wird aus (5.49) unter Ausnutzung der Additionstheoreme für trigonometrische Funktionen:

$$I_2(\vec{k}) = \underbrace{O_2(\vec{k}) \sin \{2\pi \chi(\vec{k})\}}_{\text{kohärent}} \underbrace{D_{\text{ri}}(\vec{k})}_{\text{inkoh.}} \quad (5.51)$$

mit

$$D_{\text{ri}}(\vec{k}) = \int \cos \{2\pi \vec{q} \nabla_{\vec{k}} \chi(\vec{k})\} s(\vec{q}) d\vec{q} \quad (5.52)$$

Die in Gleichung (5.51) gewonnene Produktdarstellung setzt sich aus dem bekannten kohärenten Ausdruck für die Bildintensität sowie aus einem Dämpfungsterm D_{ri} , der auf die räumlich inkohärente Quelle zurückzuführen ist, zusammen.

Betrachtet man die Dämpfungsfunktion D_{ri} in (5.52) als Funktion von $\nabla \chi(\vec{k})$, so ist der Dämpfungsterm einfach die Fouriertransformierte der Quellverteilung $s(\vec{q})$. Dies ist jedoch genau die Phasenkorrelationsfunktion $\Gamma(\vec{r}_1 - \vec{r}_2)$.

$$D_{\text{ri}}(\vec{k}) = \Gamma(\nabla \chi(\vec{k})) \quad (5.53)$$

Dieses Ergebnis ist folgendermaßen zu interpretieren: Die Steigung der Aberrationsfunktion $\chi(\vec{k})$ läßt sich durch einen Vektor $\vec{b}(\vec{k})$ darstellen, der die Dimension einer Länge hat. Der Dämpfungsterm D_{ri} ist die Phasenkorrelationsfunktion $\Gamma(\vec{r}_1 - \vec{r}_2)$ an der Stelle $\vec{r}_1 - \vec{r}_2 = \vec{b}(\vec{k})$. Daraus lassen sich einige Schlüsse ziehen:

1. Bei kohärenter Beleuchtung ist $\Gamma = \text{const.} = 1$ für jedes Argument $\vec{b}$. Dies bedeutet $D_{\text{ri}}(\vec{k}) = 1$ für jede Raumfrequenz $\vec{k}$, es findet keine Dämpfung statt.
2. Aus $\Gamma(0) = 1$ folgt, daß an allen Stellen an denen $\nabla \chi(\vec{k})$ verschwindet, $D_{\text{ri}} = 1$ ist. Dies bedeutet, daß die kohärente sowie die partiell kohärente Transferfunktion *identisch* sind in Raumfrequenzbereichen, in denen $\chi(\vec{k})$ konstant ist. An

Stellen jedoch, wo die Aberrationsfunktion sich schnell mit $\vec{k}$ ändert, also der Gradient $\vec{b}$ groß ist, werden die größten Anforderungen an die Quelle gestellt. Die Kohärenzlänge Λ muß größer sein als dieser Gradient, um eine dämpfungs-freie Kontrastübertragung zu gewährleisten.

3. Da die kohärente CTF die Aberrationsfunktion χ zum Argument hat ($B_C = \sin \chi$), sind die Stellen stärkster Dämpfung diejenigen schneller Oszillation der CTF über der Raumfrequenz k .
4. Die Oszillationen der CTF werden für große $|\vec{k}|$, also für kleine Objektdetails, immer schneller. Ab einer bestimmte Raumfrequenz k_i wird die Dämpfung abhängig vom Defokus total. Es wird keine Information mehr übertragen.
5. Den stärksten Anteil an der Steigung der Aberrationsfunktion hat der Term $C_s k^4$, der Defokussierungsterm geht nur quadratisch in k ein. Um die Steigung der Aberrationsfunktion für möglichst große Raumfrequenzen klein zu halten, ist eine möglichst kleine sphärische Aberration von entscheidender Bedeutung. Sie ist sozusagen die „treibende Kraft“ für den Anstieg der Aberrationsfunktion und somit für die totale Dämpfung hoher Raumfrequenzen verantwortlich.

Das exakte Aussehen des Dämpfungsterms D_{ri} läßt sich durch Fouriertransformation der Quellverteilung $s(|\vec{q}|)$ einfach ermitteln. Für die Zylinderverteilung erhält man:

$$D_{\text{ri}}^{\text{Zylinder}}(\vec{k}) = 2 \frac{J_1(2\pi q_0 |\nabla \chi(\vec{k})|)}{2\pi q_0 |\nabla \chi(\vec{k})|} \quad (5.54)$$

Hierbei handelt es sich bei J_1 um die Besselfunktion erster Ordnung. Für die Gaußverteilung ergibt sich als Dämpfungsterm:

$$D_{\text{ri}}^{\text{Gauß}}(\vec{k}) = \exp(-\pi^2 q_0^2 |\nabla \chi(\vec{k})|^2) \quad (5.55)$$

mit dem Gradienten der Aberrationsfunktion:

$$\nabla \chi(\vec{k}) = (C_s \lambda^2 |\vec{k}|^2 + Z) \lambda \vec{k} \quad (5.56)$$

In der letzten Gleichung wird der Defokus, von dem die Dämpfungsterme abhängen, durch die Variable Z repräsentiert. Diese Nomenklatur wird im folgenden beibehalten.

Partielle zeitliche Kohärenz

Mit dem Begriff partielle zeitliche Kohärenz trägt man der Tatsache Rechnung, daß der einfallende Elektronenstrahl nicht monochromatisch ist, sondern sich aus einem Kontinuum an Wellenvektoren zusammensetzt. Im Gegensatz zum Falle der partiellen räumlichen Kohärenz unterscheiden sich diese Wellenvektoren jedoch nicht in ihrer Richtung sondern in ihrem Betrage. Aus diesem Grunde wird die räumliche Kohärenz oft als transversale, die zeitliche als longitudinale Kohärenz bezeichnet. Ursachen für die nur partielle zeitliche Kohärenz des Elektronenstrahls sind:

1. Die Energie der durch thermische Emission aus dem Filament austretenden Elektronen gehorcht der Maxwellverteilung. Obwohl die Breite der Verteilung nur etwa 1–2 eV beträgt, spielt diese Unschärfe eine Rolle.
2. Die Beschleunigungsspannung (max. 400 kV beim JEOL 4000EX) ist aus technischen Gründen nicht genau konstant zu halten.
3. Der durch die Spulen der Objektivlinse fließende Strom unterliegt ebenfalls Schwankungen, die sich in Schwankungen der „Brekraft“ und somit der Brennweite der Linse äußern.

Da die Brennweite einer Elektronenlinse von der Energie der einfallenden Elektronen abhängt, äußern sich alle drei besprochenen Effekte in einer Streuung der Brennweite um einen bestimmten Mittelwert. Diese in der vorliegenden Arbeit Defokusstreuung (engl. *defocus spread*) genannte Größe Δ läßt sich folgendermaßen formulieren [56]:

$$\Delta = C_c \left(\frac{\Delta E}{eV} + \frac{\Delta V}{V} - \frac{2\Delta I}{I} \right) \quad (5.57)$$

Hierbei sind C_c die sogenannte *chromatische Aberrationskonstante*, ΔE die Streuung der Austrittsenergie der Elektronen am Filament, ΔV und ΔI die Schwankungen der Beschleunigungsspannung sowie des Objektivlinsenstroms um deren Mittelwerte V und I . (Das Minuszeichen vor der relativen Stromschwankung ist darin begründet, daß eine Erhöhung des Linsenstroms die Brechkraft erhöht, während eine Erhöhung der Austrittsenergie um ΔE ebenso wie die Erhöhung der Beschleunigungsspannung um ΔV sich so auswirken, als habe die Linse eine kleinere Brechkraft.) Bei dieser Formulierung der Defokusstreuung geht man davon aus, daß sich die einzelnen Unschärfen addieren.

Eine andere Möglichkeit, die Defokusstreuung zu definieren, lautet [57]:

$$\Delta = C_c \sqrt{\left(\frac{\Delta E}{eV} \right)^2 + \left(\frac{\Delta V}{V} \right)^2 + \left(\frac{2\Delta I}{I} \right)^2} \quad (5.58)$$

Bei dieser optimistischeren Formulierung geht man von einer teilweisen Kompensation der drei Effekte, die zur Defokusstreuung beitragen, aus.

Es existieren zwei verschiedene Möglichkeiten, den Effekt der partiellen zeitlichen Kohärenz mathematisch zu behandeln. Ein Weg besteht darin, die drei oben genannten Verteilungen gesondert zu behandeln. In [56] wird die Maxwellverteilung ΔE durch eine Gaußverteilung genähert, während für die Strom- und Spannungsschwankungen eine sinusförmige Verteilung angenommen wird. Der daraus resultierende Dämpfungsterm D_{zi} für zeitlich inkohärente Beleuchtung besteht dann ebenfalls aus drei Komponenten:

$$D_{zi}(\Delta E, \Delta V, \Delta I) = D(\Delta E) D(\Delta V) D(\Delta I) \quad (5.59)$$

Hierbei ist $D(\Delta E)$ eine Gaußverteilung in k , $D(\Delta V)$ und $D(\Delta I)$ sind Besselfunktionen nullter Ordnung in k^2 . Diese mathematisch „unangenehmen“ Ausdrücke werden

im allgemeinen umgangen [53, 54, 58, 55], indem die drei Effekte, die zur Defokusstreuung führen, mit einer einzigen Gaußverteilung beschrieben werden. Dieser Weg, nach dem auch die für diese Arbeit konzipierten Programme geschrieben sind, soll nun nachvollzogen werden.

Die auf 1 normierte Gaußverteilung, mit der die Defokusstreuung beschrieben wird, lautet:

$$f(Z) = \frac{1}{\sqrt{\pi}\Delta} \exp\left(-\frac{Z^2}{\Delta^2}\right) \quad (5.60)$$

Dies führt dazu, daß der Kreuzkoeffizient in Gleichung (5.38) noch zusätzlich über diese Gaußverteilung gemittelt werden muß. Der Mittelung über die verschiedenen Einfallswinkel der „Substrahlen“ im Rahmen der transversalen Kohärenz entspricht nun die Mittelung über die verschiedenen Wellenlängen dieser Strahlen zur Berechnung des Einflusses der partiellen longitudinalen Kohärenz. Man erhält für die endgültige Form des Kreuzkoeffizienten aus diesem Grunde das Doppelintegral:

$$\begin{aligned} T(\vec{k}'', \vec{k}', Z') &= \int \int s(\vec{q}) f(Z) a(\vec{q} + \vec{k}'') a(\vec{q} + \vec{k}') \\ &\times \exp\left\{-2\pi i[\chi(\vec{q} + \vec{k}'', Z + Z') - \chi(\vec{q} + \vec{k}', Z + Z')]\right\} \\ &\times d\vec{q} dZ \end{aligned} \quad (5.61)$$

Die einzige Änderung gegenüber der im Vorangehenden gezeigten Berechnung zur partiellen räumlichen Kohärenz besteht nun in dieser Erweiterung des Kreuzkoeffizienten. Mit der Annahme eines schwachen Phasenobjektes (lineare Bildberechnung + 90° Phasenverschiebung) wiederholt sich der Rechenweg. Die Aberrationsfunktion wird jetzt jedoch noch zusätzlich um den Defokus Z' in erster Ordnung entwickelt:

$$\chi(\vec{k} + \vec{q}, Z + Z') \approx \chi(\vec{k}) + \vec{q} \nabla \chi(\vec{k}) + Z \underbrace{\frac{\partial \chi(\vec{k}, Z')}{\partial Z'}}_{\chi'_Z} + \dots \quad (5.62)$$

Der entsprechende Dämpfungsterm für partiell zeitlich kohärente Beleuchtung ergibt sich ebenso wie im Falle der räumlichen Kohärenz durch Fouriertransformation der entsprechenden Verteilung – in diesem Falle $f(Z)$. Wurde im vorangehenden Falle $\nabla \chi$ als Argument der Fouriertransformierten benutzt (siehe Gl. (5.62)), muß jetzt χ'_Z eingesetzt werden. Die Fouriertransformierte der Gaußverteilung $f(Z)$ ist wiederum eine Gaußverteilung. Der Dämpfungsterm lautet somit:

$$D_{zi} = \exp\left(-\pi^2 \Delta^2 \chi'^2_Z\right) \quad (5.63)$$

mit der Ableitung

$$\chi'_Z = \frac{\lambda}{2} |\vec{k}'|^2 \quad (5.64)$$

Man beachte, daß dieser Dämpfungsterm weder vom Defokus Z noch von der sphärischen Aberration C_s abhängt. An Mikroskopparametern gehen hier – abgesehen von

Schwankungen des Objektivlinsenstroms – primär die Eigenschaften der Quelle ein. Dies sind die Wellenlänge des einfallenden Strahls, sowie deren Schwankung, die hier mit der Stromschwankung der Objektivlinse zur Defokusstreuung zusammengefaßt wurde.

Effektive Kontrasttransferfunktion

Faßt man den Einfluß partieller zeitlicher und partieller räumlicher Kohärenz zusammen, erhält man für die effektive Kontrasttransferfunktion die Produktdarstellung:

$$B_{\text{eff}}(\vec{k}) = \underbrace{\sin(\chi(\vec{k}))}_{\text{kohärent}} \underbrace{D_{\text{ri}}(\vec{k}) D_{\text{zi}}(\vec{k})}_{\text{inkohärent}} \quad (5.65)$$

Eine Darstellung der effektiven CTF des JEOL 4000EX für zwei verschiedene Defokuswerte findet sich in Abb. 5.7. An diesen Beispielen ist zu sehen, daß die kohärente CTF durch die oben besprochenen Dämpfungsterme entscheidend modifiziert wird. Speziell der totale Zusammenbruch jeglicher Kontrastübertragung bei hohen Raumfrequenzen ist von einschneidender Bedeutung. Die höchste Raumfrequenz k_i , ab der die effektive CTF – abhängig vom Defokus – bleibend verschwindet, legt das *Informationslimit* des Mikroskops fest. In [59] wurde für ein JEOL 4000EX ein gemessenes Informationslimit von 0.119 nm veröffentlicht.

Betrachtet man die Fouriertransformierte eines hochaufgelösten Bildes von ausreichend dünnem und amorphem C, Si, oder Ge, auf die die Idealisierung des schwachen Phasenobjekts zutrifft, erhält man die radiale Intensitätsverteilung der Transformierten:

$$I(k) = |O_2(k)|^2 B_C^2(k) D_{\text{ri}}^2(k) D_{\text{zi}}^2(k) = |O_2(k)|^2 B_{\text{eff}}^2(k) \quad (5.66)$$

Da die Funktion O_2 , die in diesem Falle die atomaren Streuamplituden repräsentiert, nur langsam mit k variiert, stellt die Intensitätsverteilung der Fouriertransformierten im wesentlichen das Quadrat der effektiven CTF dar.

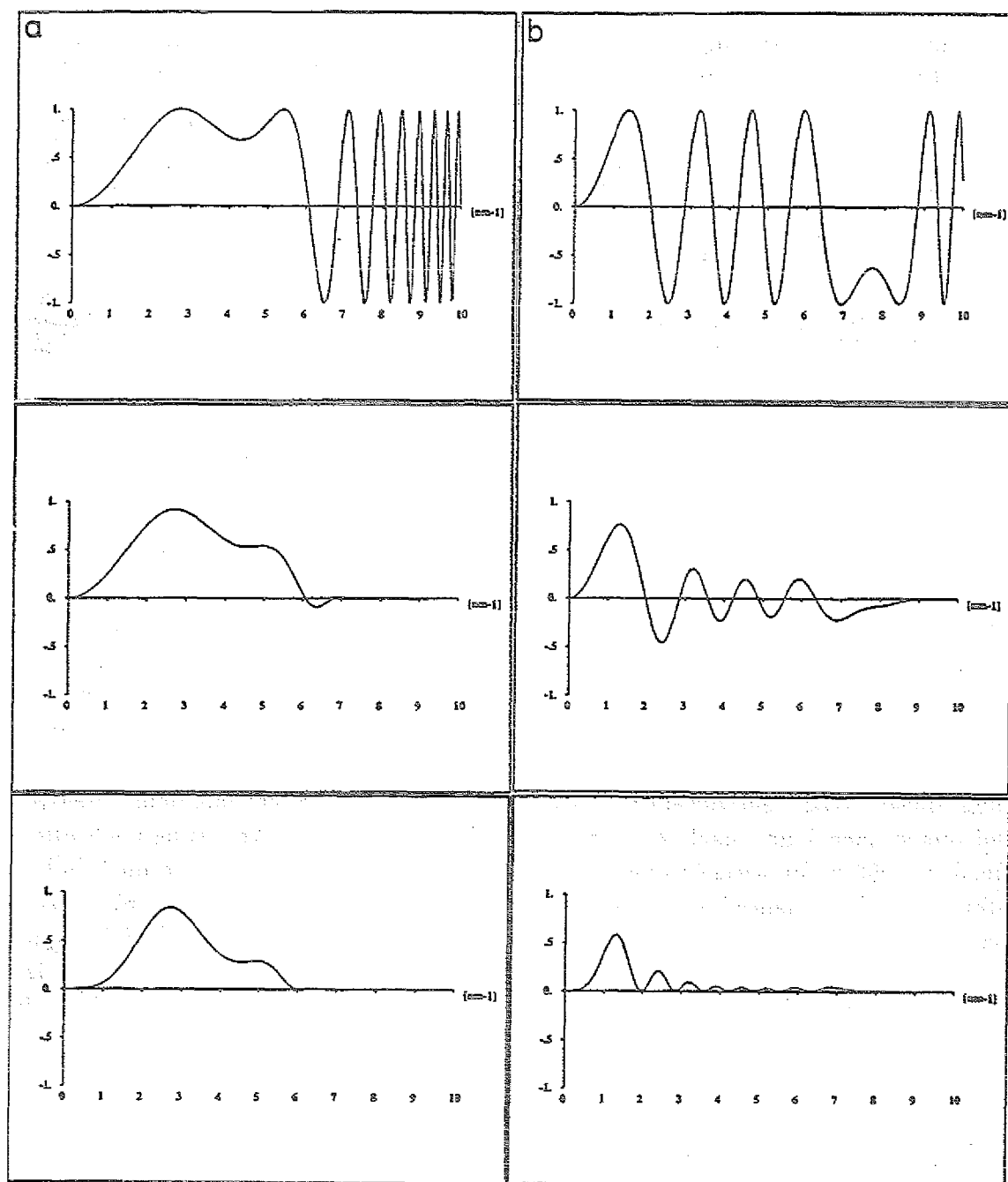


Abbildung 5.7: Oben: Kohärente Kontrasttransferfunktion; Mitte: effektive; Unten: Betragsquadrat der effektiven Kontrasttransferfunktion des JEOL 4000EX für die Defokuswerte: a) $Z = -50$ nm, b) $Z = -160$ nm. Parameter: $\Delta = 10$ nm, $\theta_c = 0.9$ mrad bei 400 kV.

5.2 Nichtlineare Kontrastübertragung

Läßt man die Idealisierung des schwachen Phasenobjekts fallen, müssen zwei Vereinfachungen, die die bisherige Darstellung geprägt haben, zurückgenommen werden:

- Die Annahme des reinen Phasenobjekts (90° -Phasenverschiebung, $O_1 = 0$) ist im allgemeinen Fall der dynamischen Beugung nicht zu rechtfertigen. Die CTF, deren Herleitung auf dieser Annahme basiert, verliert für die Beschreibung des Kontrastes eines kristallinen Objekts schon ab relativ geringen Probendicken ihre Gültigkeit (Siehe Abb. 4.5).
- Die Intensität der abgebeugten Strahlen darf im allgemeinen Fall der dynamischen Beugung nicht mehr als schwach im Verhältnis zum Nullstrahl angesehen werden. Dies impliziert, daß der im Objektspektrum nichtlineare Term $I_3(\vec{k})$ aus (5.41) mit

$$I_3(\vec{k}, Z) = \int T(\vec{k}' + \vec{k}, \vec{k}', Z) O(\vec{k}' + \vec{k}) O^*(\vec{k}') d\vec{k}' \quad (5.67)$$

ebenso nicht mehr vernachlässigt werden darf. Die volle Berechnung des Doppelintegrals (5.61) für den Kreuzkoeffizienten wird dadurch erforderlich.

Das Integral in (5.61) ist analytisch nur über eine Reihenentwicklung des Aberrationsterms im Exponenten zugänglich. Die volle Lösung hingegen ist bisher nur numerisch möglich. Die Lösungsansätze zur Berechnung des Kreuzkoeffizienten unterscheiden sich in der Ordnung bis zu der die Aberrationsfunktion entwickelt wird. Die Reihenentwicklung, die im EMS Programm IM1 (Abb. 4.4) und in den im Rahmen dieser Arbeit entwickelten Programmen benutzt wird, enthält neben den Ableitungen nach $\vec{k}$ und nach Z noch zusätzlich eine gemischte Ableitung nach diesen beiden Größen. Inzwischen existieren schon analytische Lösungen, die noch höhere Terme der Reihenentwicklung der Aberrationsfunktion einbeziehen [60, 61]. Im folgenden ist – im Gegensatz zur linearen Bildtheorie – die räumliche Beziehung der beiden Argumentvektoren des Kreuzkoeffizienten im Fourierraum von entscheidender Bedeutung. Man kann den Kreuzkoeffizienten als vierdimensionales Gebilde im Fourierraum ansehen.

Die besprochene Reihenentwicklung des Aberrationsterms aus (5.61) lautet:

$$\begin{aligned} & \chi(\vec{k} + \vec{k}'', Z + Z') - \chi(\vec{k} + \vec{k}', Z + Z') \\ &= \{ \chi(\vec{k}'', Z') - \chi(\vec{k}', Z') \} + \vec{k} (\chi''_k - \chi'_k) \\ &+ Z (\chi''_Z - \chi'_Z) + \vec{k} Z (\chi''_{kZ} - \chi'_{kZ}) + \dots \end{aligned} \quad (5.68)$$

mit den Ableitungen

$$\begin{aligned} \chi'_k &= \left. \frac{\partial \chi(\vec{k}, Z)}{\partial \vec{k}} \right|_{\vec{k}=\vec{k}', Z=Z'} = (C_s \lambda^2 |\vec{k}'|^2 + Z') \lambda \vec{k}' \\ \chi'_Z &= \left. \frac{\partial \chi(\vec{k}, Z)}{\partial Z} \right|_{\vec{k}=\vec{k}', Z=Z'} = \frac{\lambda}{2} |\vec{k}'|^2 \end{aligned}$$

$$\chi'_{kZ} = \left. \frac{\partial^2 \chi(\vec{k}, Z)}{\partial \vec{k} \partial Z} \right|_{\vec{k}=\vec{k}', Z=Z'} = \lambda \vec{k}' \quad (5.69)$$

Die Integration (5.61) liefert eine Produktdarstellung des Kreuzkoeffizienten, die – abgesehen vom Aperturterm $a(\vec{k}'') a(\vec{k}')$ – aus fünf Anteilen besteht. (Die explizite Darstellung dieser Anteile befindet sich der Übersichtlichkeit halber im Anhang B)

$$T(\vec{k}'', \vec{k}', Z) = \underbrace{a(\vec{k}'') a(\vec{k}') \exp \{ -2\pi i [\chi(\vec{k}'', Z) - \chi(\vec{k}', Z)] \}}_{\text{kohärent}} \times \underbrace{P(\vec{k}'', \vec{k}', Z) E_s(\vec{k}'', \vec{k}', Z) E_f(\vec{k}'', \vec{k}') E_x(\vec{k}'', \vec{k}')}_{\text{partiell kohärent}} \quad (5.70)$$

1. Der Exponentialterm entsteht auf leicht ersichtliche Weise bei der Betragsquadratbildung der Austrittswellenfunktion, *nachdem* die Fourierkoeffizienten $O(\vec{k})$ durch Heranmultiplizieren der durch das Mikroskop induzierten Phasenverschiebung mit $\exp(-2\pi i \chi(\vec{k}))$ beaufschlagt wurden. Betrachtet man diesen Term alleine, so stellt er den Kreuzkoeffizienten bei total kohärenter Beleuchtung dar.
2. Der komplexe Phasenfaktor P beschreibt eine zusätzliche Phasenverschiebung zum Interferenzbeitrag zweier Strahlen $\vec{k}''$ und $\vec{k}'$, die auf den Einfluß der Elektronenquelle zurückzuführen ist. Es läßt sich jedoch abschätzen [55], daß diese Phasenverschiebung von untergeordneter Bedeutung ist.
3. Der Dämpfungsterm E_s ist auf die partielle räumliche Kohärenz der Elektronenquelle zurückzuführen. Er ist vom Defokus Z abhängig und entspricht dem im Rahmen der linearen Bildtheorie behandelten Dämpfungsterm D_{ri} . Der fundamentale Unterschied zur linearen Bildtheorie besteht in der starken Abhängigkeit von der gegenseitigen räumlichen Lage der Vektoren $\vec{k}'$ und $\vec{k}''$ im Fourierraum. Diese Dämpfungsfunktion, die im Rahmen der linearen Bildtheorie im Fourierraum eindimensional ist, ist ebenso wie die anderen Beiträge zum Kreuzkoeffizienten nun in vier Dimensionen zu behandeln (k'_x, k'_y, k''_x, k''_y).
4. Der Funktion E_f entspricht in der linearen Bildtheorie der Dämpfungsterm D_{zi} . Beide sind auf die partielle zeitliche Kohärenz des Elektronenstrahls zurückzuführen und hängen selbst nicht vom Defokus ab. Auch bei E_f ist die räumliche Lage der Argumentvektoren von entscheidender Bedeutung.
5. Die Funktion E_x hat ihren Ursprung in der gemischten Ableitung bei der Reihenentwicklung der Aberrationsfunktion. Aus diesem Grunde repräsentiert sie einen Dämpfungsanteil, der auf die kombinierte Wirkung von partiell zeitlicher und partiell räumlicher Kohärenz zurückzuführen ist. Im Vergleich zu den beiden anderen Dämpfungsfunktionen E_s und E_f ist E_x – wie sich durch Abschätzung [55] zeigen läßt – von untergeordneter Bedeutung. Erst wenn extrem große Werte für die Verteilungsbreiten Δ und q_0 angenommen werden, liefert E_x einen nennenswerten Beitrag zur Gesamtdämpfung.

Der Kreuzkoeffizient $T(\vec{k}'', \vec{k}', Z)$ beinhaltet den gesamten gerätespezifischen Einfluß des Elektronenmikroskops auf den Bildkontrast. Setzt man ihn gleich 1, erhält man als Bildintensität einfach das Betragsquadrat der Austrittswellenfunktion. Da der Kreuzkoeffizient im Fourierraum definiert ist, ist seine Auswirkung im Realraum, also im Bild nicht direkt einleuchtend. Er beinhaltet die Phasen- und Amplitudenmodulation einer Bildfrequenz $|\vec{g}| = |\vec{k}'' - \vec{k}'|$, die durch Interferenz von zwei „Strahlen“ $\vec{k}''$ und $\vec{k}'$ entsteht. Da ein und dieselbe Bildfrequenz $|\vec{g}|$ meist durch Linearkombination verschiedener Paare $(\vec{k}'', \vec{k}')$ gebildet werden kann, ist über alle Möglichkeiten $(\vec{k}'', \vec{k}')$ zu summieren. Dies ist der tiefere Sinn der Autokorrelation (5.30). Jeder dieser Einzelbeiträge zur Bildfrequenz $|\vec{g}|$ wird durch den Kreuzkoeffizienten aufgrund der Lageabhängigkeit seiner Argumentvektoren, also je nach „Entstehungsgeschichte“, selektiv phasen- und amplitudenmoduliert (Siehe Anhang A).

Aufgrund der Vierdimensionalität des Kreuzkoeffizienten im Fourierraum sowie seiner Abhängigkeit vom Defokus und aufgrund der Tatsache, daß er alle Mikroskopparameter $(\lambda, C_s, \Delta, q_0)$ enthält, ist eine anschauliche und umfassende Darstellung kaum möglich. In Abb. 5.8 wurde der Versuch gemacht, einen optischen Eindruck zu vermitteln, indem das Produkt der „wichtigen“ Dämpfungsterme $E_s \times E_f$ bei festgehaltenem Defokus und festgehaltenem Vektor $\vec{k}''$ mit variablem $\vec{k}'$ dargestellt wurde. Als $\vec{k}''$ wurde der (420) Reflex von Ni_4Mo gewählt. Das in z-Richtung aufgetragene Produkt $E_s \times E_f$ gibt an, wie stark der Interferenzbeitrag des (420) Strahls mit einem in der (k'_x, k'_y) -Ebene liegenden Strahl in die Gesamtbildintensität eingeht. An der Stelle $\vec{k}' = (420)$ der k'_x, k'_y -Ebene ist $E_s \times E_f$ gleich 1, da für beliebige „Strahlen“ $\vec{k}$ gilt: $T(\vec{k}, \vec{k}, Z) = 1$. In Abbildung 5.9 ist dieselbe Darstellung von $E_s \times E_f$ für festgehaltenen (130) Strahl zu sehen.

Während der (130) Strahl im Scherzerfokus neben dem Nullstrahl mit einer größeren Anzahl anderer Strahlen nichtlinear gekoppelt ist, ist die Kontrastübertragung beim (420) Strahl sehr selektiv: Der (420) Strahl trägt auf dem Wege der linearen Übertragung überhaupt nicht zum Bildkontrast bei. Lediglich Interferenzbeiträge mit dem (330)- und dem (400) Strahl werden im Scherzerfokus nennenswert übertragen. Diese Interferenzbeiträge erzeugen dieselben Bildfrequenzen wie der (110) Strahl bzw. der (020) Strahl auf linearem Wege.

An diesem Beispiel ist deutlich zu erkennen, daß alleine die Betrachtung der linearen CTF zum Verständnis der Bildentstehung nicht ausreicht. Im Scherzerfokus ist die lineare CTF für die zum (420) Strahl gehörende Raumfrequenz bereits total gedämpft (0-Punkt in Abb. 5.8). Trotzdem trägt der (420) Strahl zum Bildkontrast bei.

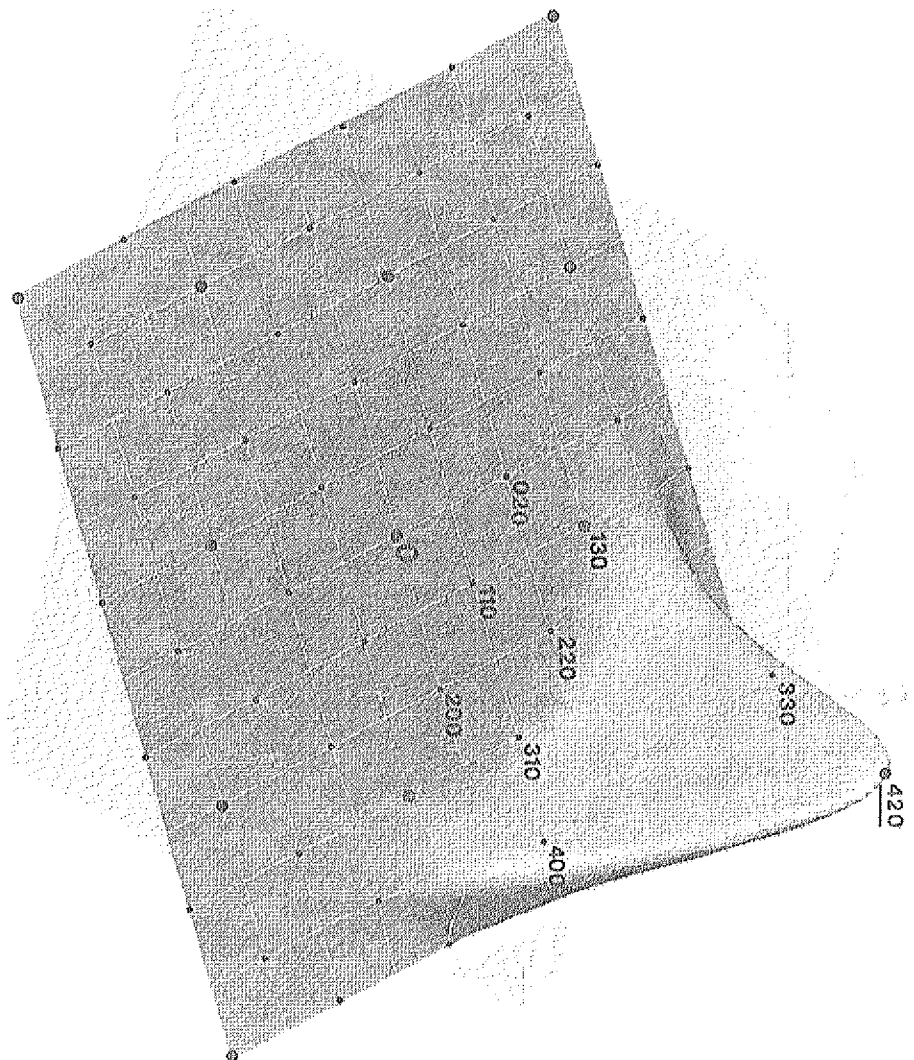


Abbildung 5.8: Darstellung des Produkts $E_x \times E_y$ (z -Achse) für festgehaltenen Strahl mit Wellenvektor $k'' = (4\ 2\ 0)$ über der k'_x, k'_y -Ebene. Die Indizierung bezieht sich auf Ni_4Mo , sodaß die nullte Lauezone der Dl_a Struktur dargestellt ist. Große Punkte zeigen den Nullstrahl bzw. die Matrixstrahlen, kleine die Überstrukturstrahlen. Parameter: $\lambda = 0.001644$ nm, $C_s = 1.0$ mm, $\Delta = 10$ nm, $q_0 = 0.55$ nm $^{-1}$, $Z = -50$ nm (Scherzerfokus).

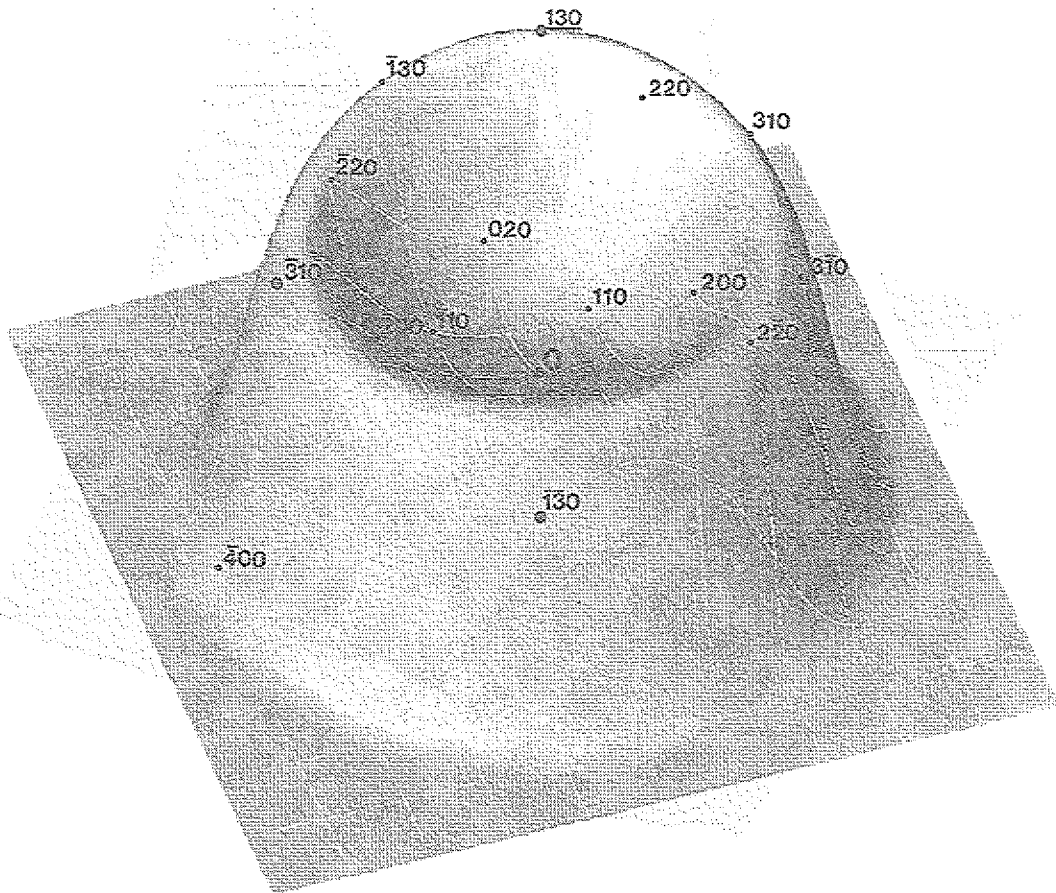


Abbildung 5.9: Darstellung des Produkts $E_s \times E_f$ (z -Achse) für festgehaltenen Strahl mit Wellenvektor $\vec{k}'' = (1\ 3\ 0)$ über der k'_x, k'_y -Ebene. Die Indizierung bezieht sich auf Ni_4Mo , sodaß die nullte Lauezone der D1_a Struktur dargestellt ist. Große Punkte zeigen den Nullstrahl bzw. die Matrixstrahlen, kleine die Überstrukturstrahlen. Parameter: $\lambda = 0.001644\text{ nm}$, $C_s = 1.0\text{ mm}$, $\Delta = 10\text{ nm}$, $q_0 = 0.55\text{ nm}^{-1}$, $Z = -50\text{ nm}$ (Scherzerfokus).

Kapitel 6

Experiment

6.1 Präparation von Ni₄Mo Proben

Das für diese Arbeit benutzte Probenmaterial aus ferngeordnetem Ni₄Mo wurde freundlicherweise von Herrn Dr. Mayer, Max-Planck-Institut für Metallforschung, Stuttgart, in Form eines 180 µm dicken Probenstreifens zur Verfügung gestellt.

Die Präparation der elektronenmikroskopischen Proben erfolgte durch elektrolytisches Polieren. Hierfür wurden aus dem Probenstreifen Scheibchen mit einem Durchmesser von 3 mm gestanzt. Beim Polieren im Düsenstrahlverfahren wird von beiden Seiten ein Säurestrahl auf eine ca. 1 mm² große Fläche des Scheibchens gespritzt. Das elektrolytische Polieren der Scheibchen erfolgte in einem Bad aus 50 ml HClO₄ und 950 ml CH₃COOH. Dabei wurde an die als Anode geschaltete Probe eine Polierspannung von 90 V angelegt. Die Temperatur des Säurebads betrug 10° C. Nach dem muldenförmigen Abdünnen des Probenscheibchens durch den Säurestrahl bildete sich nach 1 – 2 Minuten ein Loch von etwa 0.1 mm Durchmesser. Am Rand dieses Lochs, an dem die Dicke des Materials in Form eines flachen Keils ansteigt, ist die Probe für einen Elektronenstrahl durchstrahlbar. Der Poliervorgang wurde mittels eines Lichtstrahls, der durch das Loch im Probenscheibchen auf eine Photozelle fällt, automatisch abgebrochen. Die Probe wurde durch vorsichtiges Eintauchen in Methanol und durch ein anschließendes kurzes Ethanolbad von Säurerückständen gereinigt und sofort in einem Exsikkator gelagert.

Im Schnitt war nur eine von vier Proben für den Einsatz am Hochauflösungsmikroskop JEOL 4000EX tauglich. Dies lag daran, daß das Probenmaterial polykristallin war und der durchstrahlbare Randbereich jeweils von nur wenigen zufällig orientierten Körnern gebildet wurde. Um eine Abbildung entlang der [001] Zonenachse machen zu können, wären in den meisten Fällen Probenkippen nötig gewesen, die mit dem Goniometer des JEOL 4000EX, das nur eine maximale Kippung von ±15° erlaubt, nicht erreichbar waren.

Es stellte sich heraus, daß nur „frisch“ präparierte Proben für eine hochauflösende Untersuchung tauglich waren. Bei Proben, die mehrere Tage alt waren, ließ der Bildkontrast vermutlich aufgrund sich an der Oberfläche ausbildender Oxidschichten merklich nach.

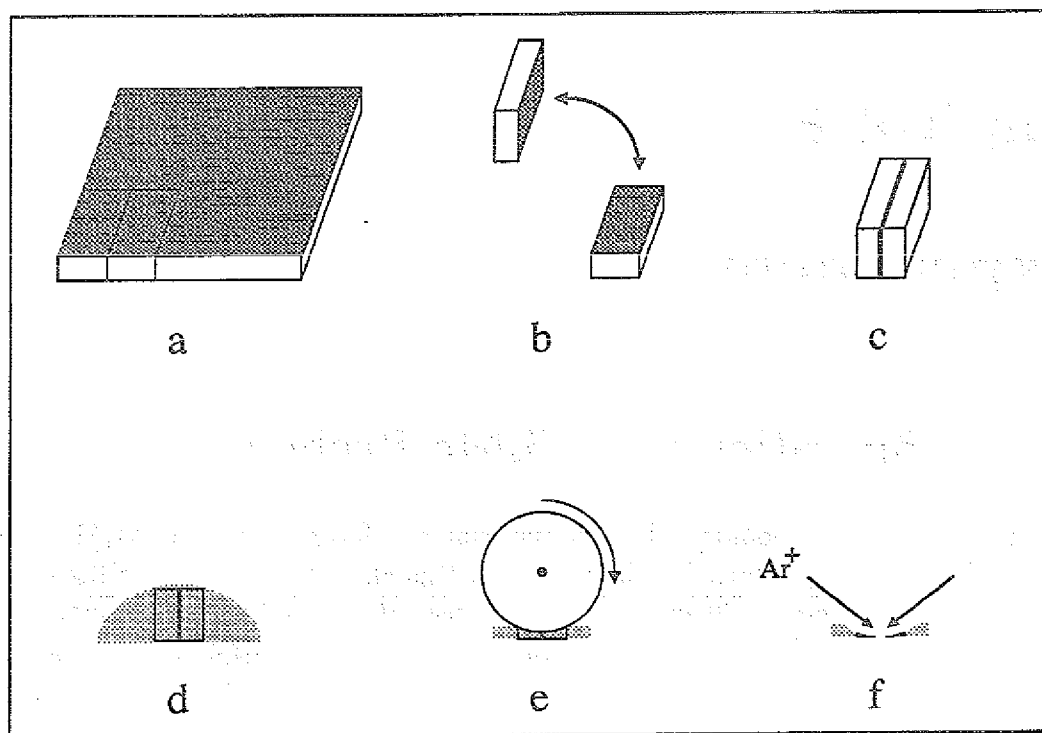


Abbildung 6.1: Schematische Darstellung der Präparation von Querschnittsproben.

6.2 Präparation von $\text{YBa}_2\text{Cu}_3\text{O}_7$ Querschnittsproben

Die in dieser Arbeit gezeigten Hochauflösungsaufnahmen von $\text{YBa}_2\text{Cu}_3\text{O}_7$ wurden an einer etwa 200 nm dünnen Schicht aufgenommen, die mittels *off-axis sputtern* epitaktisch auf einem Substrat aus SrTiO_3 deponiert worden war. Bei der untersuchten Schicht handelt es sich um einen sogenannten c-Achsen Film. Dies bedeutet, daß die c-Achse von $\text{YBa}_2\text{Cu}_3\text{O}_7$ in Richtung der Substratnormalen orientiert ist, die ihrerseits wiederum parallel einer [100] Richtung von SrTiO_3 ausgerichtet ist. Die Probe in Form eines 10×10 mm großen Chips wurde im Institut für Schicht- und Ionentechnik an der KFA Jülich hergestellt und von Fr. Dipl.-Phys. R. Dömel freundlicherweise zur Verfügung gestellt.

Da man bei Aufsicht auf die Schichtoberfläche parallel zur c-Achse von $\text{YBa}_2\text{Cu}_3\text{O}_7$ blickt, mußten für Abbildungen entlang der [100] Zonenachse sogenannte Querschnittsproben angefertigt werden. Die einzelnen Bearbeitungsschritte sind in Abbildung 6.1 schematisch dargestellt.

Aus dem $10 \text{ mm} \times 10 \text{ mm}$ großen Substrat wurden Stäbchen der Abmessung $1 \text{ mm} \times 2.5 \text{ mm}$ mittels einer Diamantsäge (Marke Accutom, Sägeblatt $d=0.15 \text{ mm}$) herausgeschnitten (Abb. 6.1a). Jeweils zwei dieser Stäbchen wurden an der Schicht-

seite mittels Zweikomponentenkleber (Araldit Kleber und IPD Härter) zusammengeklebt (Abbildungen 6.1b und 6.1c). Die Aushärtung des Klebers erfolgte in einem Ofen bei 110°C. Anschließend wurden die so zusammengeklebten Stäbchen in einen Tropfen desselben Klebers gebettet, der wiederum im Ofen ausgehärtet wurde (Abb. 6.1d). Die Einbettung dient zur mechanischen Unterstützung des äußerst brüchigen Substratmaterials. Mittels einer Rasierklinge wurde vom Rand her überschüssiger Kleber weggeschnitten, sodaß die Proben einen Durchmesser von 3 mm haben und in den vorgesehen Platz des Probenhalters des Mikroskops passen.

Danach wurden die Proben vorsichtig manuell mittels SiC Schleifpapier auf eine Restdicke von etwa 100 μm heruntergeschliffen. Die Unterseite des verbliebenen Plättchens wurde danach mit Aluminiumoxid-Suspension glattpoliert. Die Oberseite wurde mit einem Schleifrad muldenförmig bis auf eine Restdicke von etwa 10 μm oder weniger abgedünnt (Abb. 6.1e) und anschließend ebenfalls mittels Aluminiumoxid-Suspension glattpoliert. Nach diesem Vorgang wurden die Proben auf einem Kupferferring von 3 mm Durchmesser aufgeklebt.

Die Proben wurde dann in einer Ionendünnungsanlage der Firma Gatan auf einem rotierbaren, mit flüssigem Stickstoff gekühlten Halter befestigt und mit 5 kV Ar^+ Ionen beschossen. Hierbei wurden pro Stunde etwa 3 μm Material abgetragen sodaß je nach Anfangsdicke nach einigen Stunden in der Mitte der Probe ein Loch entstand (Abb. 6.1f). Sobald das Loch mit einem Lichtmikroskop wahrnehmbar war, wurde der Vorgang unterbrochen. Anschließend erfolgte in derselben Anlage eine kurze Nachbehandlung der Oberfläche durch Beschuß mit 3 kV Ar^+ Ionen, die unter einem Winkel von 12° eingestrahlt wurden. Ebenso wie bei dem im vorangehenden Abschnitt beschriebenen elektrolytischen Polierverfahren ist der aus diesem Verfahren resultierende flach ansteigende Lochrand mit Elektronen durchstrahlbar.

6.3 Mikroskopie

In Abb. 6.2 ist das im Rahmen dieser Arbeit am Institut für Mikrostrukturforschung benutzte 400 kV Mikroskop JEOL 4000EX dargestellt. Die Mikroskopsäule steht zur Vermeidung von Vibrationen von der Konsole getrennt auf einem ca. 60 Tonnen schweren Betonblock. Dieser ist seinerseits ölgedämpft auf Federn gelagert. Am 4000EX ist eine Videokamera installiert, die zum einen zur direkten Bildbeobachtung mit einem Monitor verbunden ist. Die Videokamera speist zudem noch ein digitales Bildverarbeitungssystem der Fa. Tietz, Gauting. Dieses System ermöglicht neben einer Vielfalt von Bild- und Kontrastmanipulationen eine „on-line“ Fouriertransformation. Die Rechenzeit für die Transformation eines Bildes aus 512×512 Bildpunkten beträgt etwa eine Sekunde.

Neben den gerätetechnischen Voraussetzungen entscheidet die bei der Justage des Mikroskops aufgewandte Sorgfalt darüber, ob man die gewünschte Auflösung letztendlich erreicht. Diese Justage ist am Anfang jeder Sitzung zu erneuern.

Neben dem Zentrieren der Kondensorblende und der Korrektur des Kondensorastigmatismus wird hierbei mittels Ablenkspulen der Elektronenstrahl über die gesamte Säulenlänge möglichst geradlinig entlang der optischen Achse des Instru-

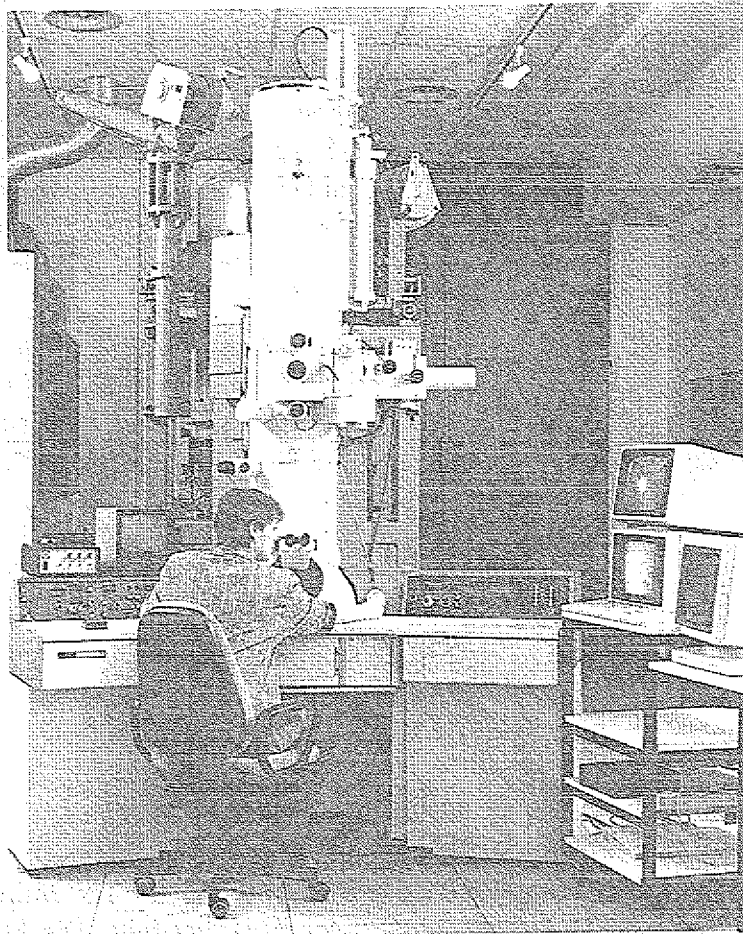


Abbildung 6.2: Das JEOL 4000EX des Instituts für Mikrostrukturforschung.

ments geführt. Für die Hochauflösung sind alle Justageschritte, die die Objektivlinse betreffen, von besonderer Bedeutung. Folgende Punkte verdienen besondere Beachtung:

Anregung der Objektivlinse

Die sphärische Aberrationskonstante C_s hängt von der Linsenanregung ab. Trägt man die berechnete sphärische Aberration über der Anregung auf, so hat die Kurve ein flaches Minimum bei einem bestimmten Objektivstrom. Dieser Wert ist vom Hersteller spezifiziert und wurde schon zu Anfang einer jeden Sitzung eingestellt. Da die Defokussierung sich gleichermaßen aus Linsenbrennweite und dem Abstand Probenaustrittsfläche – Objektivlinse zusammensetzt, muß bei der weiteren Arbeit soweit wie möglich mittels der Höhenverstellung des Probenhalters mechanisch fokussiert werden. Neben der Gewährleistung minimaler sphärischer Aberration ist die Tatsache von Bedeutung, daß die im folgenden beschriebene Korrektur von Astigmatismus und Koma nur für einen schmalen Bereich der Linsenanregung durchzuführen ist. Entfernt man sich im Laufe der Arbeit von der bei der Justage benutzten Anregung zu weit, so sind die anfänglich gemachten Korrekturen hinfällig.

Korrektur des Objektivastigmatismus

Die Objektivlinse ist aufgrund mangelnder Rotationssymmetrie ihres Magnetfeldes ohne spezielle Korrektur meist astigmatisch. Dies bedeutet, daß die Brennweite der Linse entlang zweier senkrecht zueinander stehender Achsen verschieden ist. Der Astigmatismus macht sich in der fehlenden Zentrosymmetrie des Diffraktogramms bzw. der Fouriertransformierten von einem hochaufgelösten Bild eines amorphen Objekts bemerkbar. Zur Korrektur dient der Rand der Probe, der aufgrund der Probenpräparation meist nicht kristallin ist. Der Astigmatismus läßt sich unter Beobachtung der „on-line“ Fouriertransformierten mit Hilfe spezieller Ablenkspulen, die in der Nähe der Objektivlinse angebracht sind, korrigieren. Er muß bei starken Änderungen der Linsenanregung sowie bei Wechseln der Objektivblende nachkorrigiert werden.

Korrektur der Koma

Die theoretische Auflösung eines Mikroskops läßt sich nur bei idealer geometrischer Abstimmung des strahlerzeugenden Systems und des Magnetfeldes der Objektivlinse erreichen. Dies setzt folgende Bedingungen voraus [62]:

- Das Zentrum des Magnetfelds in der Objektebene wird in das Zentrum der Bildebene abgebildet.
- Der Elektronenstrahl fällt parallel zur Symmetrieachse des Linsenfelds ein.
- Das Magnetfeld der Objektivlinse ist zentrosymmetrisch.

Die erste Bedingung läßt sich durch präzise mechanische Anordnung von Objektivlinse, Zwischenlinsen und Projektorlinse entlang einer Achse fest einstellen. Die zweite Bedingung ist durch eine Kippung des Elektronenstrahls parallel zur so festgelegten optischen Achse erfüllbar. Jedoch ist das Magnetfeld der Objektivlinse im allgemeinen nicht zentrosymmetrisch. Dieser Linsenfehler, der selbst bei geometrisch richtiger Justage eine Richtungsabhängigkeit der Aberrationsfunktion bewirkt, heißt axiale Koma. Die axiale Koma wirkt sich so aus, als würde man – eine ideale Linse vorausgesetzt – mit gekipptem Strahl arbeiten. Sie läßt sich durch eine geeignete Kippung des Strahls in die sogenannte komafreie Achse kompensieren. Hierzu wurde nach Ausführung aller sonst erforderlichen Justageschritte folgende Prozedur durchgeführt [63]:

Die ungefähre Lage der komafreien Achse wurde durch periodische Variation der Hochspannung mittels des sogenannten „*high tension wobblers*“ ermittelt. Bei Variation der Beschleunigungsspannung rotiert das Bild um das sogenannte Spannungszentrum. Der Strahl wurde mittels Ablenkspulen so gekippt, daß das Spannungszentrum mit der Bildschirmmitte übereinstimmte. Bei 500 000 -facher Vergrößerung wurde anschließend der Strahl zu beiden Seiten dieser neu gefundenen Achse mittels des „*tilt wobblers*“ periodisch hin und her gekippt. Dabei wurde das Bild des amorphen Probenrandes auf dem Monitor betrachtet. Die Strahlkippfung induziert genauso wie die axiale Koma ihrerseits wieder Astigmatismus sowie eine Änderung des Defokus. Weicht die „Mittelstellung“, um die der Strahl dabei gekippt wird, von der komafreien Achse ab, so addieren bzw. subtrahieren sich je nach Kipprichtung die Effekte von Koma und Kippung zu zwei verschiedenen Bildern. Durch geeignete Korrektur der anfangs durch Justage des Spannungszentrums gefundenen „Mittelstellung“, um die der Strahl periodisch gekippt wird, können die Bilder angeglichen werden.

Die für jegliche Art von quantitativer Elektronenmikroskopie wünschenswerte Anfertigung von Defokus-Serien erwies sich am JEOL 4000EX als äußerst schwieriges Unterfangen. Grund hierfür waren mechanische Vibrationen, die beim Wechsel der Filmplatten zwischen zwei Aufnahmen ausgelöst wurden. Die Vibrationen, die sich auf die Mikroskopsäule und die Probenhalterung ausdehnten, machten die Anfertigung einer qualitativ hochwertigen Defokus-Serie zu einem seltenen Zufallsergebnis.

Zum einen verstellte sich der Defokus durch eine Positionsänderung der Probe parallel zur optischen Achse. Die Defokusänderungen waren teilweise so massiv, daß nach einer photographischen Aufnahme kein Bildkontrast mehr auf dem Monitor zu beobachten war, was auf Höhenverstellungen von einigen -zig Nanometern schließen läßt. Ebenso wurde die Probe in der zur optischen Achse senkrechten Objektebene ausgelenkt. Bis die dadurch ausgelöste Probendrift zum Erliegen kam, verstrichen unnötig lange Wartezeiten zwischen den Aufnahmen. Der dritte und bei weitem unangenehmste Effekt der mechanischen Erschütterungen war die Änderung der Probenorientierung zwischen zwei Aufnahmen. Die Orientierung einer Probe mittels der Kippvorrichtung des Probenhalters muß vor einer Aufnahme mit einer Genauigkeit, die im Milliradbereich liegt, erfolgen. Die durch die Vibrationen ausgelösten Verkippungen waren jedoch so stark, daß es meist nicht gelang, dieselbe Probenstelle

zweimal hintereinander in korrekter Orientierung aufzunehmen.

Zur Simulation des elektronenmikroskopischen Bildkontrasts müssen die elektronenoptischen Parameter möglichst genau bekannt sein. Die im Rahmen dieser Arbeit gemachten experimentellen Aufnahmen und Simulationsrechnungen basieren auf einer Beschleunigungsspannung von 400 kV. Bei den Simulationsrechnungen wurde für die sphärische Aberrationskonstante C_s die Herstellerangabe von $C_s = 1 \text{ mm}$ übernommen. Die Werte für die Defokusstreuung Δ und den Semikonvergenzwinkel der Beleuchtung θ_c wurden vom Verfasser im Rahmen seiner Diplomarbeit mittels der Frankschen Methode gemessen [58]. Dabei wurde für die Defokusstreuung ein Wert von $\Delta = (10 \pm 2) \text{ nm}$ und für den Semikonvergenzwinkel ein Wert von $\theta_c = (0.9 \pm 0.2) \text{ mrad}$ ermittelt. Diese Werte stimmen sehr gut mit Literaturwerten überein, die an JEOL 4000EX Mikroskopen mit unterschiedlichen Methoden ermittelt wurden [59, 64].

Kapitel 7

Quantitativer Bildvergleich

7.1 Einführung

Wie in den vorangehenden Kapiteln ausführlich erläutert wurde, ist der Bildkontrast hochauflöster Aufnahmen nicht allein von der projizierten Struktur des Objekts abhängig. Insbesondere die Dicke des durchstrahlten Objektbereichs und die Fokussierung der Objektivlinse beeinflussen nachhaltig den Bildkontrast. Bildsimulationen, mittels derer man aus dem Grad der Übereinstimmung mit experimentellen Bildern bestimmte Strukturmodelle ausschließen oder verifizieren kann, erfordern jedoch gerade die Eingabe der meist unbekannten Parameter Defokus und Probenstärke. Bisherige quantitative Verfahren zur Ermittlung dieser Parameter werden im folgenden kurz aufgeführt.

Ein oft verwendetes Verfahren zur Bestimmung des Defokus besteht in der optischen Beugung eines Laserstrahls an einem hochauflösten Negativ, auf dem eine amorphe Probenstelle abgebildet ist [65]. Zum einen erfordert die Methode, die auf der Bestimmung von Nullstellen der Kontrasttransferfunktion (siehe Gl. (5.66) sowie Abb. 5.7) in einem optischen Diffraktogramm beruht, daß die beobachtete amorphe Stelle der Näherung des schwachen Phasenobjekts genügt. Die Methode kann deshalb nur in äußerst dünnen Probenbereichen angewandt werden, die Atome niedriger Kernladungszahl beinhalten. Wenn weder die Dicke der Probenstelle noch die chemische Zusammensetzung genau bekannt sind, ist es schwierig abzuschätzen, ob die Näherung des schwachen Phasenobjekts noch gültig oder bereits unzulässig ist. Zudem ist es nötig, daß sich die amorphe Stelle in der Nähe des zu untersuchenden Probenbereichs befindet, da sich der Defokus aufgrund einer möglichen Neigung der Probenunterseite lokal ändern kann. Eine zusätzliche Schwierigkeit dieses Verfahrens besteht darin, daß bei Überfokussierung oder bei Fokussierung nahe der Scherzbedingung nur ein einzelnes breites Transferband in der Kontrasttransferfunktion auftritt. Die Methode, die ja auf der Bestimmung von Nullstellen beruht, ist hier wegen eines Mangels an Meßpunkten äußerst ungenau.

Eine andere Möglichkeit, die Fokussierung zu bestimmen, bietet sich direkt im Experiment: Stellt man die Anregung der Objektivlinse so ein, daß sich an einer amorphen Probenstelle minimaler Kontrast ergibt, kann man für diese Stelle den Defokus nach Gl. (5.25) ermitteln. Der gewünschte Wert kann dann durch zusätzli-

ches Über- oder Unterfokussieren eingestellt werden. Diese Methode setzt ebenso wie die zuvor besprochene Möglichkeit voraus, daß die abgebildete Stelle als schwaches Phasenobjekt beschrieben werden kann. Es ergeben sich somit ähnliche Nachteile. Zudem setzt die Methode voraus, daß die Defokussierung zwischen der Einstellung des Kontrastminimums und der Aufnahme keinen Schwankungen unterliegt. Dies ist jedoch häufig infolge der durch die Aufheizung der Probe eintretenden Probendrift, die eine Änderung des Abstandes Probenunterseite – Bildebene bewirkt, nicht der Fall.

Ähnliche Probleme treten bei der lokalen Bestimmung der durchstrahlten Probendicke auf. Die Bestimmung der Probendicke mittels konvergenter Elektronenbeugung ist nur für Probendicken von einigen hundert Nanometern geeignet [66]. In der hochauflösenden Elektronenmikroskopie haben die durchstrahlten Probenbereiche jedoch typischerweise eine Dicke unter 20 nm.

Der Gebrauch von Proben, die sich wohldefiniert entlang hochsymmetrischer kristallographischer Ebenen keilförmig spalten lassen, eröffnet die Möglichkeit der lokalen Dickenbestimmung [67]. Bei bekanntem Keilwinkel kann die Dicke durch Messung des Abstands zwischen der beobachteten Probenstelle und dem Keilrand bestimmt werden. Allerdings ist die Anwendung dieser Methode in den seltensten Fällen möglich, da sich nur wenige Materialien wie z.B. Elementhalbleiter oder III-V Verbindungshalbleiter entlang geeigneter Ebenen spalten lassen.

Zusätzlich zu den oben erwähnten direkten Meßmethoden eröffnet das Abzählen von Pendellösungsoszillationen des ungebeugten Strahls eine Abschätzung der Probendicke (Siehe Kapitel 4). In diesem Falle muß die Austrittswellenfunktion zuvor berechnet werden, um die Extinktionslänge zu ermitteln. Bei nur langsamer Variation der Amplitude des Nullstrahls mit der durchlaufenen Probenstrecke ist eine genaue Dickenbestimmung jedoch schwierig.

Eine häufig benutzte Alternative zu den oben erwähnten Methoden der Defokus- und Dickenbestimmung besteht in der Simulation einer großen Anzahl verschiedener Bilder ein und derselben bekannten Probenstelle, die sich ausschließlich durch die Parameter Defokus und Probendicke unterscheiden. Die experimentellen Parameter werden dann durch die beste Übereinstimmung zwischen Simulation und Experiment bestimmt. Die Übereinstimmung wird hierbei per „Augenschein“ festgestellt. Die Nachteile einer solchen Vorgehensweise sind augenfällig: Da das Maß an Übereinstimmung nur qualitativ beurteilt wird, spielt die subjektive Meinung des Betrachters eine Rolle. Außerdem ist diese Methode sehr zeitaufwendig, da bei Materialien mit komplizierter Mikrostruktur kleine Änderungen von Defokus oder Probendicke wesentliche Änderungen des Bildkontrasts bewirken und somit eine große Anzahl von Bildern berechnet und verglichen werden muß.

Um die Nachteile der zuletzt genannten Methode zu eliminieren, wurde im Rahmen dieser Arbeit ein Programmpaket für den schnellen Vergleich zwischen experimentellen und simulierten Bildern entwickelt. Das Paket dient dazu, einen bekannten und periodisch aufgebauten Bereich einer experimentellen Aufnahme automatisch mit allen simulierten Bildern zu vergleichen, die durch Variation von Defokus und Probendicke entstehen können. Die dabei eingesetzten schnellen Berechnungs- und Vergleichsmethoden ermöglichen ein lückenloses „Absuchen“ der gesamten Defokus-

Dicken Parameterebene. Durch Auftragung der Kreuzkovarianzkoeffizienten zwischen einem experimentellen Bild und einer großen Anzahl simulierter Bilder (typischerweise 10^4) in dieser Ebene können die Parameter, die den experimentellen Bildkontrast reproduzieren, auf einen Blick erfaßt werden.

7.2 Theorie

Der Vergleich zwischen einem experimentellen Eingabebild und der Gesamtheit der durch Variation von Defokus und Probendicke möglichen Simulationsbilder wird im Fourierraum durchgeführt. Der Bildvergleich im Fourierraum bietet folgende Vorteile:

Die Bildberechnung wird normalerweise in zwei Stufen durchgeführt [55]: Zuerst wird die Fouriertransformierte der Bildintensität $I(\vec{k})$ ermittelt, danach erfolgt die Transformation der Bildintensität in den Realraum. Der zweite Schritt kostet verhältnismäßig viel Rechenzeit und ist für den Bildvergleich nicht notwendig, da die gesamte Bildinformation auch in der Fouriertransformierten $I(\vec{k})$ gespeichert ist.

Ein anderer Vorteil der Darstellung der Bildintensität im Fourierraum liegt in der Tatsache begründet, daß sich im Falle typischer periodisch aufgebauter Objekte die Bildintensität durch ein Satz von wenigen δ -artigen Funktionen beschreiben läßt. Jede dieser δ -artigen Funktionen beschreibt eine Raumfrequenz und ist durch eine komplexe Zahl charakterisiert. Im Gegensatz dazu ist die Bildinformation im Realraum kontinuierlich über das gesamte Bild verteilt. Im Realraum besteht ein relativ „kleines“ Bild schon aus $64 \times 64 = 4096$ Bildpunkten, die alle berücksichtigt werden müssen. Geht man von einem vollständig periodischen Bild aus, so reduziert sich der zum Bildvergleich notwendige Rechenaufwand im Fourierraum abermals erheblich.

Ein dritter Vorteil, der sich durch die Wahl der Fourierraumdarstellung ergibt, besteht in der Möglichkeit der Rauschfilterung der Bilder. Setzt man voraus, daß nur periodische Anteile im Bild enthalten sein dürfen, erreicht man durch Weglassen aller Fourierkoeffizienten, die nicht zu δ -förmigen Maxima „kondensiert“ sind, eine beinahe perfekte Rauschfilterung. Diese Vorgehensweise liefert jedoch verfälschte Ergebnisse, wenn nicht spezielle Eigenschaften der diskreten Fouriertransformation berücksichtigt werden. Deshalb wird im folgenden großes Augenmerk darauf gerichtet, die Auswirkungen der Rauschfilterung zu überprüfen.

7.2.1 Bildsimulation

Das Hauptprogramm des zum quantitativen Bildvergleich entwickelten Programmpakets, das die simulierten Bilder berechnet und mit den experimentellen Eingabebildern vergleicht, benutzt als Ausgangsdaten die Fourierkoeffizienten $F(\vec{k})$ der Austrittswellenfunktion $\Psi(\vec{r})$:

$$\Psi(\vec{r}) = \sum_{\vec{k}} F(\vec{k}) e^{-2\pi i \vec{k} \vec{r}}. \quad (7.1)$$

Die komplexen Fourierkoeffizienten $F(\vec{k})$ können mit Hilfe des EMS-Programmpakets [46] durch Verwendung des Multislice-Algorithmus oder des Blochwellenverfahrens berechnet werden (Siehe Kapitel 4).

Die Bildintensität $I(\vec{k})$, die man im Fourierraum im Falle partiell kohärenter Einstrahlung erhält, wurde in Kapitel 5 anhand der Gleichungen (5.41), (5.61) und (5.70) erläutert und sei an dieser Stelle der Übersichtlichkeit halber noch einmal kurz dargestellt:

$$I(\vec{k}, Z) = \sum_{\vec{g}} F(\vec{g} + \vec{k}) F^*(\vec{g}) T(\vec{g} + \vec{k}, \vec{g}, Z). \quad (7.2)$$

Die zur Bildberechnung benötigten Transmissions-Kreuzkoeffizienten T werden hierbei in der Näherung von Ishizuka [55] berechnet (Siehe Anhang B).

7.2.2 Bilderkennung

In diesem Abschnitt wird gezeigt, daß die Mustererkennung mittels Berechnung des Kreuzkorrelationskoeffizienten zwischen zwei Bildern im Realraum ein perfektes Analogon im Fourierraum besitzt. Die Betrachtung geht von der Intensität $I_{\text{exp}}(\vec{r}_i)$ des experimentellen Bildes sowie der Intensität $I_{\text{sim}}(\vec{r}_i)$ des simulierten Bildes aus, die jeweils auf einem diskreten Gitter mit den Gitterpunkten $\vec{r}_i$ erfaßt werden. Die Messung der Bildintensität an Bildpunkten $\vec{r}_i$ entspricht der digitalen Aufnahme von Bildern mittels CCD- (*charge coupled device*) Kameras [68, 6, 69], wobei der Vektor $\vec{r}_i$ die Koordinate des i -ten Bildpunktes angibt. Die Intensitätswerte der einzelnen Bildpunkte können nun in beliebiger aber gleicher Reihenfolge in Vektoren $\vec{e}$ und $\vec{s}$ angeordnet werden, wobei der Buchstabe „e“ für das experimentelle und der Buchstabe „s“ für das simulierte Bild steht:

$$\vec{e} = (e_1, e_2, \dots, e_N) \quad \text{mit} \quad e_i = I_{\text{exp}}(\vec{r}_i) - \bar{I}_{\text{exp}}, \quad (7.3a)$$

$$\vec{s} = (s_1, s_2, \dots, s_N) \quad \text{mit} \quad s_i = I_{\text{sim}}(\vec{r}_i) - \bar{I}_{\text{sim}}. \quad (7.3b)$$

Da die Mittelwerte $\bar{I}_{\text{exp}}$ und $\bar{I}_{\text{sim}}$ experimenteller bzw. simulierter Bilder erfahrungsgemäß die bei weitem größten Fourierkoeffizienten sind und kaum Information tragen, wurden sie in den Gleichungen (7.3a,b) subtrahiert. Auf diese Weise kann verhindert werden, daß geringe *relative* Abweichungen der Mittelwerte zweier Bilder die Übereinstimmung der anderen Fourierkoeffizienten, die ja das eigentliche Bildmuster repräsentieren, völlig „verdecken“. Mit der Annahme, daß der Verschiebungsvektor $\vec{R}$ zwischen dem experimentellen und dem simulierten Bild null ist, lautet der normierte Kreuzkovarianzkoeffizient [70, 71]:

$$c(\vec{R} = 0) = \frac{\vec{e} \cdot \vec{s}}{|\vec{e}| |\vec{s}|}. \quad (7.4)$$

Der Koeffizient (7.4) ist der Cosinus des Winkels Φ zwischen den Vektoren $\vec{e}$ und $\vec{s}$. Ist der Kreuzkovarianzkoeffizient $c = 1$ bzw. $\Phi = 0^\circ$ so bedeutet dies, daß das experimentelle Bild und das simulierte Bild abgesehen von den jeweiligen Mittelwerten identisch sind. $c = 0$ bzw. $\Phi = 90^\circ$ bedeutet, daß die beiden Bilder unkorreliert sind.

Messungen des Winkels $\Phi = \arccos(c)$ wurden benutzt, um die lokale chemische Zusammensetzung an GaAs/AlGaAs Grenzflächen aus Hochauflösungsaufnahmen zu bestimmen [72, 73].

Im Fourierraum kann die Fouriertransformierte der Bildintensität, die im folgenden mit Großbuchstaben gekennzeichnet ist, ähnlich wie im Realraum in Vektorform angeordnet werden:

$$\vec{E} = (E_1, \dots, E_N) \quad \text{mit} \quad E_j = \frac{1}{N} \sum_i e_i e^{2\pi i \vec{k}_j \vec{r}_i}, \quad (7.5a)$$

$$\vec{S} = (S_1, \dots, S_N) \quad \text{mit} \quad S_j = \frac{1}{N} \sum_i s_i e^{2\pi i \vec{k}_j \vec{r}_i}. \quad (7.5b)$$

Im allgemeinen sind die Fourierkoeffizienten $\vec{E}_j$ und $\vec{S}_j$ komplexe Zahlen. Durch Einsetzen der Inversion der Gleichungen (7.5a,b),

$$e_i = \sum_j E_j e^{-2\pi i \vec{k}_j \vec{r}_i}, \quad (7.6a)$$

$$s_i = \sum_j S_j e^{-2\pi i \vec{k}_j \vec{r}_i}, \quad (7.6b)$$

in den Zähler des Kreuzkovarianzkoeffizienten (7.4) erhält man:

$$\vec{e} \cdot \vec{s} = N \vec{E} \cdot \vec{S}^*. \quad (7.7)$$

Von dieser Gleichung ausgehend kann man nach dem Gesetz von Parseval [70] eine Fourierdarstellung des Nenners des Kreuzkovarianzkoeffizienten (7.4) finden:

$$|\vec{e}| = \sqrt{\vec{e} \cdot \vec{e}} = \sqrt{N \vec{E} \cdot \vec{E}^*} = \sqrt{N} |\vec{E}| \quad (7.8a)$$

$$|\vec{s}| = \sqrt{\vec{s} \cdot \vec{s}} = \sqrt{N \vec{S} \cdot \vec{S}^*} = \sqrt{N} |\vec{S}| \quad (7.8b)$$

Schließlich läßt sich der Kreuzkovarianzkoeffizient in seiner Fourierdarstellung völlig analog zur Realraumdarstellung (7.4) beschreiben:

$$c(\vec{R} = 0) = \frac{\vec{E} \cdot \vec{S}^*}{|\vec{E}| |\vec{S}|}. \quad (7.9)$$

Der Bildkontrast ist durch die Größen $|\vec{e}|/\sqrt{N}$ und $|\vec{s}|/\sqrt{N}$ bzw. durch $|\vec{E}|$ und $|\vec{S}|$ definiert. Bei der Bildung des Kreuzkovarianzkoeffizienten wird der absolute Bildkontrast durch Normierung der Intensitätsvektoren $\vec{e}$ und $\vec{s}$ bzw. $\vec{E}$ und $\vec{S}$ zu Einheitsvektoren eliminiert. Dies ist zur Mustererkennung unbedingt notwendig, da sonst ein zu einem gegebenen Motiv schlecht passendes Bild aufgrund hoher Intensität den Vorzug vor einem besser passenden oder gar identischen Bild mit geringer Intensität bekäme.

7.2.3 Fourierfilterung

Das in experimentellen Bildern unvermeidlich auftretende Rauschen beeinflusst auf unerwünschte Weise den Wert der Kreuzkovarianz, die ja in Bezug zu idealen simulierten Bildern berechnet wird. Es ist möglich, das Rauschen durch Anwendung einer Filtermaske im Fourierraum zu reduzieren. Die Anwendung von Filtermasken auf nicht-periodische Bilddetails ist sehr schwierig, da sie zu Bildverzerrungen führen kann [74, 75, 76, 77, 78]. Auch die Filterung periodischer Bilder erfordert große Sorgfalt und wird deshalb ausführlich erläutert.

Im Falle eines perfekt periodischen Bildmotivs sollte im Fourierraum die gesamte Information in Form von δ -ähnlichen Maxima „kondensieren“. Die Anwendung einer Filtermaske, die ausschließlich diese Maxima „durchläßt“ entspricht einer beinahe perfekten Rauschfilterung. Nur der Anteil des Rauschens, der zufällig dieselbe Periodizität wie das Bildmotiv besitzt, kann durch diese Methode nicht entfernt werden. Betrachtet man eine diskrete Fouriertransformation, so entspricht diese Technik der Anwendung eines Bragg- Filters bei dem die Öffnungen, die an den Positionen der δ -ähnlichen Maxima liegen, genau den Durchmesser eines einzigen Bildpunktes haben.

Die auf diskrete Weise berechnete Fouriertransformierte eines perfekt periodischen und rauschfreien Bildes weicht jedoch im allgemeinen schon erheblich von der beschriebenen Form der δ -ähnlichen Maxima ab. Die Information von periodischen Bildern erstreckt sich fast kontinuierlich über große Teile des Fourierraums und eine Filtertechnik, die mit engen Öffnungen um die lokalen Maxima der Fouriertransformierten arbeitet, liefert in einem solchen Falle falsche Ergebnisse. Die Verteilung von periodischer Information im Fourierraum, die im Englischen manchmal als „leakage“ [79] bezeichnet wird, liegt an der Wahl der Begrenzung des Bildmotivs vor der Transformation. Auf diesen Effekt und seine mögliche Vermeidung wird im folgenden eingegangen.

Ausgangspunkt einer zweidimensionalen Fouriertransformation sei eine Intensitätsverteilung I_0 , die aus einem Satz diskreter Sinus- und Cosinuswellen besteht:

$$I_0(x, y) = \sum_{i,j} A_{ij} e^{-2\pi i(u_i x + v_j y)}. \quad (7.10)$$

Das die Funktion $I_0(x, y)$ begrenzende Fenster im Realraum besitze Einheitsfläche und sei definiert durch:

$$w(x, y) = \begin{cases} 1 & \text{für } -1/2 \leq x, y < 1/2 \\ 0 & \text{sonst} \end{cases} \quad (7.11)$$

Die Begrenzung des Bildmotivs durch das Fenster $w(x, y)$ verändert die ideale Fouriertransformierte $I_0(u, v)$ einer unendlich ausgedehnten Funktion nun zu

$$I_w(u, v) = I_0(u, v) \otimes W(u, v) = \sum_{i,j} A_{ij} W(u - u_i, v - v_j), \quad (7.12)$$

wobei $W(u, v)$ die Fouriertransformierte des quadratischen Fensters $w(x, y)$ ist und $\otimes$ eine Faltungsoperation darstellt. Die Faltung mit der Funktion

$$W(u, v) = \frac{\sin \pi u}{\pi u} \frac{\sin \pi v}{\pi v} = \text{sinc}(\pi u) \cdot \text{sinc}(\pi v) \quad (7.13)$$

führt zu einem „Ausschmieren“ der δ -artigen Maxima, die man im Fourierraum bei der Transformation eines unendlich ausgedehnten Motivs erhalten würde. Im Fall einer diskreten Berechnung der Fouriertransformation mit Hilfe des Cooley–Tukey Algorithmus [80] wird die Funktion $I_w(u, v)$ nur an Werten von u und v ermittelt, die ganze Zahlen sind. Dies entspricht einer Reduktion eines kontinuierlichen Satzes von Wellenvektoren (u, v) zu einem diskreten Satz (m, n) , der ausschließlich Wellen, die in Bezug auf die Abmessungen des Fensters $w(x, y)$ periodisch sind, beinhaltet. Die Periodizität des Bildmotivs spielt hierbei keine Rolle. Die ganzzahligen Werte (m, n) entsprechen zweidimensionalen Miller Indizes, wobei die Fensterfunktion $w(x, y)$ die entsprechende Einheitszelle definiert:

$$I_w(m, n) = \sum_{i,j} A_{ij} W(m - u_i, n - v_j) \quad m, n = 0, \pm 1, \pm 2, \dots \quad (7.14)$$

Falls nun das Bildmotiv so gewählt wird, daß es mit der vorgegebenen Fensterfunktion kommensurabel ist, reicht der zur Verfügung stehende reduzierte Satz von Wellenvektoren (m, n) – ähnlich den Miller Indizes – aus, um das Bildmotiv zu beschreiben. Durch Erfüllung der Kommensurabilität zwischen Bildmotiv und Fenster wird erreicht, daß nur ganzzahlige Werte (i, j) der Wellenvektoren (u_i, v_j) in der Fouriertransformierten vorkommen. In diesem speziellen Fall lautet Gleichung (7.14)

$$I_w(m, n) = \sum_{i,j} A_{ij} W(m - i, n - j) \quad i, j = 0, \pm 1, \pm 2, \dots \quad (7.15)$$

und folglich können nur ganzzahlige Vielfache von π im Argument der sinc-Funktionen auftreten. Da

$$\text{sinc}(\pi k) = \begin{cases} 1 & \text{für } k = 0 \\ 0 & \text{für } k = \pm 1, \pm 2, \dots \end{cases} \quad (7.16)$$

gilt, sind die sinc-Funktionen auf dem zweidimensionalen Pixelgitter so positioniert, daß außer deren Zentrum ausschließlich Nullstellen auf Gitterpunkten liegen. Sie verhalten sich somit wie ideale zweidimensionale δ -Funktionen:

$$I_w(m, n) = \sum_{i,j} A_{ij} \delta(m - i, n - j) = A_{mn}. \quad (7.17)$$

Indem man also genau eine ganze Zahl von quadratisch begrenzten Bildmotiven parallel zu den Kanten eines bei der numerischen Fouriertransformation üblichen quadratischen Pixelgitters anordnet, erreicht man, daß ein Fourierkoeffizient A_{mn} nur ein einziges Pixel mit den Koordinaten (m, n) auf dem entsprechenden Gitter des Fourierraums „anregt“. Auf diese Art und Weise erreicht man, daß Amplitude und Phase der Fourierkoeffizienten des Bildes ausgelesen werden können. Diese Methode eignet sich nicht nur für quadratische Bildmotive. Bilder von rechteckigen Einheitszellen und Bilder von Einheitszellen, deren Winkel von 90° abweichen, können vor der Fouriertransformation numerisch auf quadratische Form verzerrt werden. Sogar Bilder mit hexagonaler Einheitszelle können durch die Wahl einer größeren orthorhombischen Einheitszelle so digitalisiert werden, daß ein „Ausschmieren“ der Fourierkoeffizienten vermieden werden kann.

Es ist jedoch unwahrscheinlich, daß in einem experimentellen Bild benachbarte Einheitszellen genau identischen Kontrast aufweisen. Kleine Verschiebungen von Intensitätsmaxima oder Minima können zum Beispiel von lokalen Verzerrungen des Materials herrühren oder auf nicht-kristalline Schichten zurückzuführen sein, die auf den Oberflächen des Präparats haften [81, 82]. In diesem Fall ist jedoch die Bedingung strenger Periodizität nicht mehr gegeben und die Intensitätsverteilung innerhalb einer Einheitszelle kann nicht mehr in einen endlichen Satz von Sinus- und Cosinuswellen zerlegt werden. Als Folge dieser nicht-harmonischen Beiträge zum Bild ist wiederum ein „Ausschmieren“ der Fourierkoeffizienten möglich. Zusätzlich können je nach Genauigkeit des Digitalisierungsgerätes sogenannte Quantisierungsfehler den gleichen Effekt hervorrufen.

Ähnlich wie im Realraum können diese nicht-periodischen Anteile durch Mittelwertbildung über mehrere Einheitszellen reduziert werden, da kleine geometrische Verzerrungen und Intensitätsabweichungen kein langreichweitiges Verhalten aufweisen. Quantisierungsfehler können durch Mittelwertbildung ebenso reduziert werden, da zusätzliches Rauschen, das ja immer vorhanden ist, diese an sich systematischen Fehler dem Zufall unterwirft [79]. Das Einbeziehen von N Einheitszellen in den zu transformierenden Bildausschnitt vergrößert gegenüber einer einzelnen Einheitszelle das Verhältnis von Signal zu Rauschen der extrahierten Fourierkoeffizienten um einen Faktor $\sqrt{N}$. Die maximale Zahl von Einheitszellen, die in den Bildausschnitt einbezogen werden kann, ist durch die Anzahl der zur Verfügung stehenden Bildpunkte begrenzt. Diese Grenze ist durch das Whittaker–Shannon Abtasttheorem festgelegt, wonach zur Auflösung einer bestimmten Raumwellenlänge wenigstens zwei Abtastpunkte pro Periode benötigt werden [83].

Die Rücktransformation der auf diese Weise extrahierten Fourierkoeffizienten erzeugt aus einem Bild, das ursprünglich aus N verschiedenen Einheitszellen bestand, ein Bild in dem alle Einheitszellen exakt identisch sind [84]. Diese Methode kann also nur für periodische Objekte angewandt werden. Im Falle einer Anwendung auf nicht-periodische Bildinhalte liefert die Methode falsche Ergebnisse, da die intrinsische Mittelwertbildung alle nicht-periodischen Bildinhalte eliminiert anstatt sie zu verstärken [85].

7.2.4 Symmetriekorrektur

Details experimenteller Hochauflösungsbilder weichen oftmals von der idealen Symmetrie der beobachteten Struktur ab. Mögliche Ursachen für diese Fehlerquelle sind zum Beispiel eine Fehlorientierung des Objekts, eine leichte Kippung des Elektronenstrahls aus der optischen Achse oder ein nicht korrigierter Astigmatismus der Objektlinse. Diese experimentellen Fehler bewirken eine Abweichung der Amplituden und Phasen der Fourierkoeffizienten eines hochaufgelösten Bildes von den durch die Symmetrie des Objekts vorgegebenen Werten.

Obwohl die Symmetrie des Objekts durch eine Fehljustage des Elektronenstrahls gebrochen werden kann, bleibt die Periodizität der *einzelnen* Bildfrequenzen erhalten [62, 63]. Sogar eine leichte Kippung des Objekts aus der Symmetrieachse, wie sie oft zwischen Objektgebieten, die durch eine Grenzfläche getrennt sind, beobachtet

werden kann, beeinflußt nicht die Bildfrequenzen.

Vor einem Vergleich mit idealen simulierten Bildern ist es von Vorteil, die möglicherweise gebrochene Symmetrie experimenteller Bilder wiederherzustellen. Diese Prozedur ist nicht notwendig, falls die experimentellen Bilder ein gewisses Minimum an Qualität erfüllen. Bei starkem Symmetriebruch kann jedoch die Mustererkennung mittels des Kreuzkovarianzkoeffizienten (Gl. (7.4)) zu paradoxen Ergebnissen führen. Wie in Abbildung 7.1 gezeigt ist, werden Bilddetails, die gegenüber ihrer idealen Position etwas verschoben sind, durch den Kreuzkovarianzkoeffizienten unterbewertet oder gar nicht registriert. Kommen im experimentellen Bild Abweichungen von der durch die Objektstruktur vorgegebenen Symmetrie vor, empfiehlt es sich, das Wissen um die Symmetrieeigenschaften zu nutzen und durch Kompensation des oben beschriebenen Effekts die Mustererkennung robuster zu machen.

Die hier benutzte Methode zur Symmetriekorrektur ähnelt der sogenannten „Kristallographischen Bildverarbeitungsmethode“ [86, 87, 88, 89]. Diese Methode ermöglicht eine präzise Bestimmung von Atomkoordinaten aus elektronenmikroskopischen Aufnahmen von Kristallen mit sehr großen Einheitszellen, wie zum Beispiel $K_7Nb_{15}W_{13}O_{80}$. Aus einer größeren Anzahl von Symmetriegruppen, die mit dieser Methode behandelt werden können, sei die vierzählige Symmetrie zur Erläuterung herausgegriffen.

Da die Bildintensität $I(x, y)$ eine reelle Größe ist, ist die Fouriertransformierte $I(u, v)$ hermitesch, d.h. $I(-u, -v) = I^*(u, v)$. Aufgrund dieser Eigenschaft muß nur die Hälfte der Fourierkoeffizienten eines Bildes zum Bildvergleich herangezogen werden. Im Falle eines zentrosymmetrischen Bildes müssen die Imaginärteile aller Fourierkoeffizienten verschwinden falls ein Symmetriezentrum als Ursprung für die Fouriertransformation gewählt wird. In der Amplituden-Phasen Notation

$$I(u, v) = A(u, v) e^{i\varphi(u, v)}, \quad (7.18)$$

können die Phasenwinkel φ deshalb nur die Werte 0 oder π annehmen. Ein Symmetriezentrum (x_0, y_0) eines zentrosymmetrischen Bildes kann durch die Berechnung des gewichteten Phasenresiduums ermittelt werden [87]. Das gewichtete Phasenresiduum α ist definiert durch:

$$\alpha(x, y) = \frac{\sum_i A_i \cdot (\varphi_i \bmod \pi)}{\sum_i A_i} \quad (7.19)$$

Die Summe erstreckt sich über alle Fourierkoeffizienten, die sich auf einen beliebigen Koordinatenursprung (x, y) beziehen. Bei der Wahl eines Symmetriezentrums (x_0, y_0) als Ursprung ist das Phasenresiduum im Falle eines idealen zweizähligen Bildes exakt null. Aufgrund symmetriebrechender Abweichungen in experimentellen Bildern ist dies in der Realität nicht der Fall. Abweichungen vom Wert Null ergeben sich dadurch, daß nicht alle Fourierkomponenten an ein und demselben Punkt des Bildes gerade ein Maximum oder Minimum haben, wie es für ein Symmetriezentrum erforderlich wäre. Die den Fourierkomponenten entsprechenden Bildwellen sind gegeneinander verschoben und das Minimum des gewichteten Phasenresiduums bildet den besten Kompromiß für ein Symmetriezentrum. Die Abweichung des Phasenresiduums vom Idealwert Null ist deshalb ein gutes Maß für die Qualität eines Bildes.

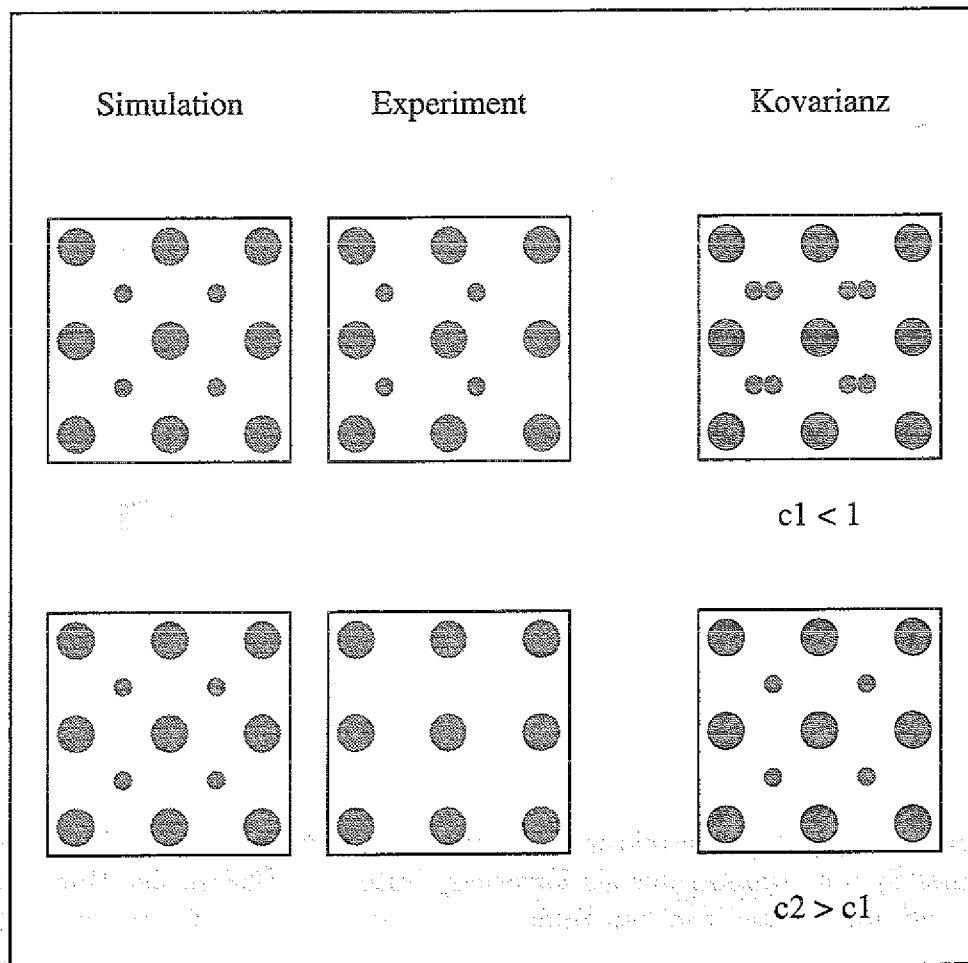


Abbildung 7.1: Paradoxes Ergebnis der Mustererkennung mittels des Kreuzkovarianzkoeffizienten. Linke Spalte: Symbolisches „simuliertes“ Bild mit perfekter Symmetrie. Mittlere Spalte: Symbolbilder, die experimentellen Bildern entsprechen sollen. Beim oberen „experimentellen“ Bild sind die kleineren Punkte gegenüber der symmetrischen Position verschoben, beim unteren Bild fehlen diese ganz. Rechte Spalte: Überlappende Gebiete, die zum Zähler des Kreuzkovarianzkoeffizienten in Gl. (7.4) beitragen, sind schwarz gekennzeichnet und für beide „experimentellen“ Bilder von gleicher Größe. Da das obere „experimentelle“ Bild einen größeren Kontrast als das untere aufweist, ist der Kreuzkovarianzkoeffizient $c1$ des oberen Bildes aufgrund der Kontrastnormierung durch den Nenner in Gl. (7.4) kleiner als der Koeffizient $c2$ des unteren Bildes. Die Kreuzkovarianz „ignoriert“ in diesem Falle Bilddetails, die verschoben sind, komplett.

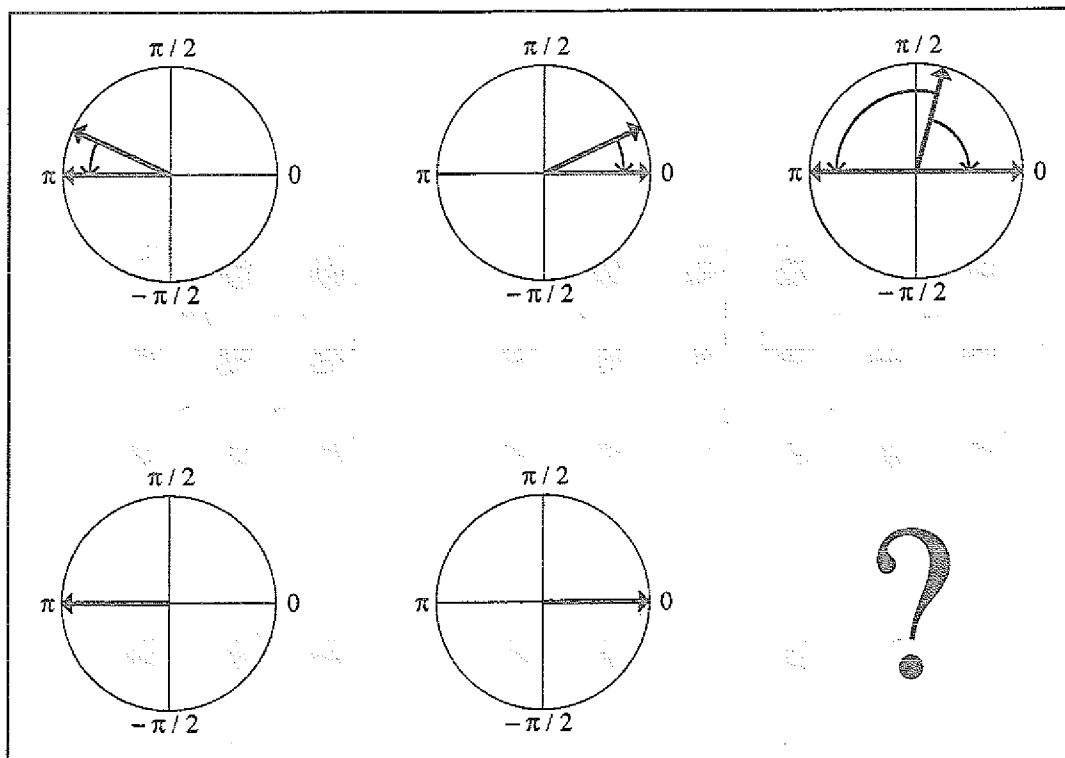


Abbildung 7.2: Symmetriekorrektur annähernd zentrosymmetrischer Bildmotive. Bei Wahl eines Symmetriezentrums als Ursprung dürfen die Phasen der Fourierkoeffizienten eines idealen zweizähligen Bildes nur die Werte π oder 0 annehmen. Obere Reihe: Unkorrigierte Fourierkoeffizienten in der komplexen Zahlenebene. Untere Reihe: Korrigierte Fourierkoeffizienten. Bei moderaten Abweichungen von den idealen Phasenlagen π oder 0 (links und mitte) ist eine Korrektur möglich. Bei starken Abweichungen vom Idealwert (rechts) ist nicht mehr zu erkennen, ob 0 oder π der „richtige“ Wert ist und eine Korrektur ist nicht mehr möglich.

Korrigierte Phasenwinkel Φ können auf einfache Weise durch Anwendung des folgenden Algorithmus gefunden werden (Siehe Abb. 7.2):

$$\Phi = \begin{cases} 0 & \text{für } -\pi/2 < \varphi \leq \pi/2, \\ \pi & \text{sonst} \end{cases} \quad (7.20)$$

Der Algorithmus ist nur dann erfolgreich anwendbar, falls die Abweichungen von den Idealwerten 0 oder π moderat sind. Je mehr sich die experimentellen Phasenwinkel den Werten $\pm\pi/2$ annähern, desto schwieriger wird es zu entscheiden, in welche Richtung die Korrektur vorgenommen werden soll und umso wahrscheinlicher wird es, daß der Algorithmus (7.20) versagt.

Im Falle vierzähliger Symmetrie müssen noch zusätzlich folgende Bedingungen

für symmetrieverwandte Fourierkoeffizienten gelten:

$$A(u, v) = A(-v, u) = A, \quad (7.21a)$$

$$\varphi(u, v) = \varphi(-v, u) = \varphi. \quad (7.21b)$$

Für den Fall, daß unterschiedliche Amplituden in einer Gruppe symmetrieverwandter Koeffizienten auftreten, kann durch simples Mitteln über $A(u, v)$ und $A(-v, u)$ die Vierzähligkeit wiederhergestellt werden. Im korrigierten Bild wird der resultierende Mittelwert allen vier Koeffizienten einer Gruppe zugewiesen.

Wie bereits erwähnt wurde, scheitert die Phasenkorrektur, falls Phasenverschiebungen nahe $\pi/2$ in experimentellen Bildern auftreten. Die Wahrscheinlichkeit, den richtigen Phasenwinkel zu finden, ist in solch einem Falle nur 1:1. Um falsche Phasenkorrekturen zu vermeiden, werden solche „problematischen“ Koeffizienten in der hier besprochenen Prozedur nicht verändert. Fourierkoeffizienten werden nur dann korrigiert, falls folgende Bedingung erfüllt ist:

$$\cos(\varphi(u, v)) \cdot \cos(\varphi(-v, u)) > 0. \quad (7.22)$$

Dies bedeutet, daß zwei symmetrieverwandte Koeffizienten in der selben komplexen Halbebene liegen müssen und entweder $\text{Re}(I(u, v)), \text{Re}(I(-v, u)) > 0$ oder $\text{Re}(I(u, v)), \text{Re}(I(-v, u)) < 0$ gelten muß. Als Folge dieser zusätzlichen Bedingung können auch symmetriekorrigierte Bilder noch schwache Symmetrieabweichungen aufweisen. Solche Bilder im werden folgenden „partiell symmetriekorrigierte Bilder“ genannt.

7.3 Bilderfassung und Bildverarbeitung

Die hochaufgelösten Bilder wurden auf Kodak SO-163 Filmplatten aufgenommen, die acht Minuten in einer D19 Lösung entwickelt wurden. Kleine Ausschnitte der Negative wurden mittels einer CCD- (*charge coupled device*) Kamera, die über ein Balgengerät mit einem 200 mm Zoom-Objektiv verbunden war, auf einem Feld von 512×512 Bildpunkten digital erfaßt. Bei maximaler Vergrößerung dieses optischen Aufbaus beträgt die digitalisierte Fläche des Negativs 2×2 mm, was einer Auflösung von $4 \times 4 \mu\text{m}$ pro Bildpunkt entspricht. Bei einer primären Vergrößerung des Elektronenmikroskops von 500 000 entspricht die Distanz zwischen zwei Bildpunkten einer Objektdistanz von 8 pm. Nach dem Whittaker-Shannon Abtasttheorem [83] ist somit Objektinformation von wenigstens 20 pm Auflösung auf der digitalisierten Aufnahme gespeichert. Da in diesem Falle der granulare Kontrast des Negativs, der hauptsächlich durch das statistische Rauschen der aufgenommenen Elektronen verursacht wird, schon aufgelöst wird und da die elektronenmikroskopisch auflösbaren Objektdetails eine minimale Abmessung von 100 pm besitzen, ist eine solch hohe optische Nachvergrößerung nicht sinnvoll. Die Bilder, die in diesem Kapitel gezeigt werden, wurden deshalb mit einer 2–3 fach geringeren optischen Vergrößerung aufgenommen um einen möglichst großen Bildausschnitt zu erfassen. Diese Vergrößerung reicht aus, um eine Objektperiodizität von 0.06 nm aufzulösen. Dieser Wert entspricht der Hälfte des Informationslimits des JEOL 4000EX, das bei 0.12 nm liegt [59].

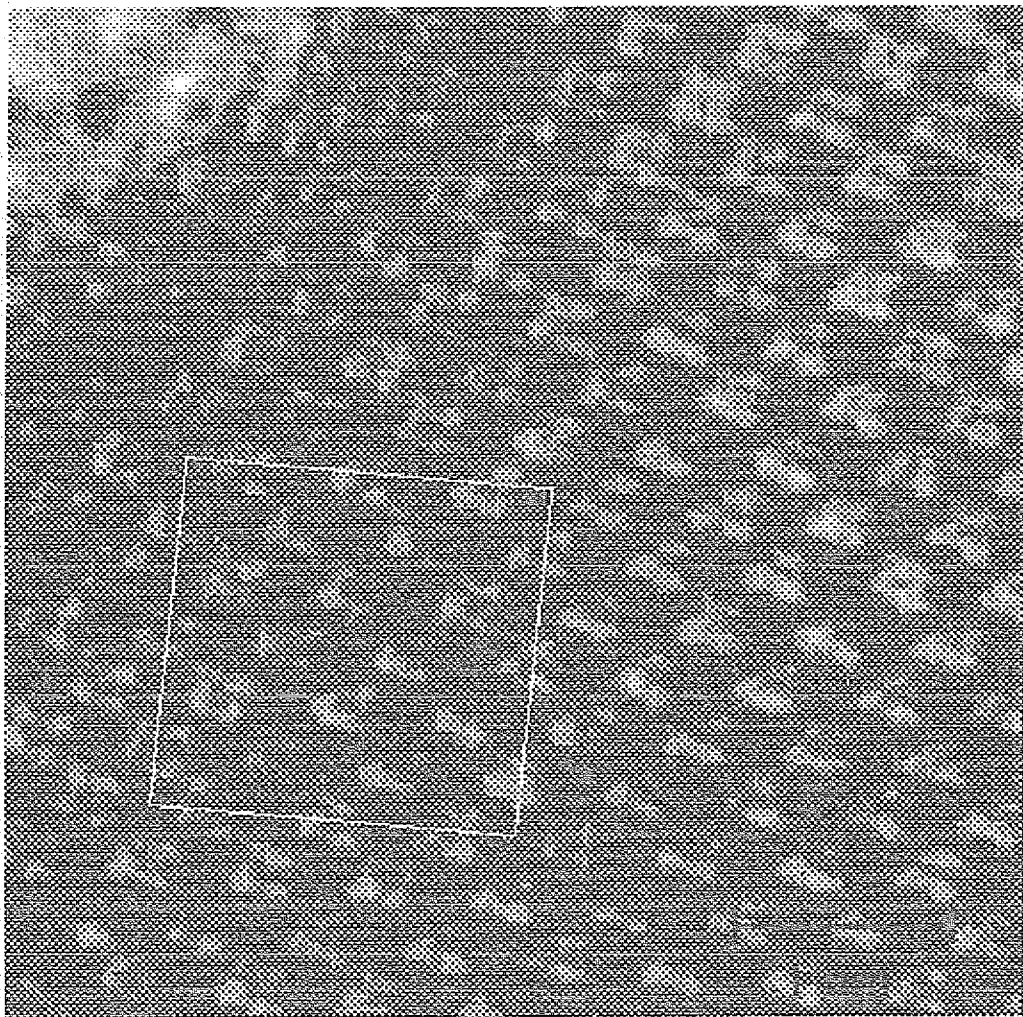


Abbildung 7.3: Hochaufgelöstes Bild von Ni₄Mo, das parallel zur [001] Kristallachse aufgenommen wurde. Die Abbildung besitzt eine Auflösung von 512×512 Bildpunkten und zeigt die erste Stufe der Digitalisierung des Negativs. Der eingezeichnete Rahmen zeigt den Ausschnitt innerhalb dessen eine zweite Digitalisierung auf ein Feld von 256×256 Bildpunkten erfolgte. Wie bei allen nachfolgenden Aufnahmen in diesem Kapitel wurde zur Darstellung der volle Bereich einer 256-stufigen Grauskala benutzt.

Die Bilddaten, die auf einem PC mittels einer „frame grabber“ Karte zwischengespeichert wurden, wurden zur Weiterverarbeitung zu einer VAX station 3100 transferiert. In einer ersten Bearbeitungsstufe erfolgte die Korrektur der nichtlinearen Intensitätsempfindlichkeit des gesamten Aufnahmesystems, das die photographische Filmplatte und die CCD- Kamera umfaßt. Die Reaktion der CCD- Kamera auf Objekte unterschiedlicher optischer Dichte wurde mit Hilfe eines geeichten halbtransparenten Graustufenkeils der Firma Kodak für alle Blendeneinstellungen des Zoom-Objektivs ermittelt. Um die Sensitivität der Filmplatten gegenüber der einfallenden Elektronenintensität zu messen, wurde eine Serie homogen ausgeleuchteter Negative mit ansteigender Belichtungszeit am Mikroskop aufgenommen. Der jedem Negativ entsprechende Wert der optischen Dichte wurde mit Hilfe der CCD- Kamera ermittelt. Hierbei wurde festgestellt, daß die auf den Negativen gemessene optische Dichte bis zu etwa einem Wert von 1.8 linear mit der Belichtungszeit ansteigt. Dies deckt sich gut mit dem in [90] festgestellten Verhalten des Kodak SO- 163 Films. Eine spezielle Routine, die die gemessenen Werte optischer Dichte für linear ansteigende Belichtungszeiten als auch die zuvor gemessene Reaktion der CCD- Kamera auf verschiedene Werte optischer Dichte in Form von Tabellen enthält, wird zur automatischen Konversion CCD-Daten in eine lineare Skala der einfallenden Elektronenintensität eingesetzt.

Photographische Filme besitzen eine sogenannte Modulationstransferfunktion M , die die Intensität hoher Raumfrequenzen auf dem Film stärker dämpft als die niedriger Raumfrequenzen. Die Modulationstransferfunktion besitzt die Form $M(g) = 1/(1 + g^2 x_0^2)$ wobei g die infragekommene Raumfrequenz auf dem Film bezeichnet und x_0 eine filmspezifische charakteristische Länge ist. Beim Kodak 4463 Film tritt eine 50 %-ige Dämpfung der Intensität bei einer charakteristischen Raumfrequenz $g_0 = 1/x_0$ von etwa 40 mm^{-1} auf [91]. Dieser Wert für g_0 wurde zur Korrektur der Intensitätswerte des hier verwendeten Kodak SO-163 Film übernommen. Die folgende Abschätzung zeigt, daß der Einfluß der Modulationstransferfunktion auf die Bildintensität bei den hier gewählten Aufnahmebedingungen eher geringfügig ist. Eine der größten Raumfrequenzen, die im Ni_4Mo System für eine Abbildung mit dem JEOL 4000EX überhaupt noch in Frage kommt, hat der (420) Strahl, der mit $g = 7.8 \text{ nm}^{-1}$ nahe am Informationslimit liegt. Dem entspricht nach 500 000 -facher Vergrößerung auf dem Negativ eine Raumfrequenz von $g = 15.6 \text{ mm}^{-1}$. Für diese Frequenz hat die Modulationstransferfunktion bei der angenommenen charakteristischen Raumfrequenz von $g_0 = 40 \text{ mm}^{-1}$ einen Wert von 0.87. Der bei der Messung der Fourierkoeffizienten im *schlimmsten* Fall durch die Vernachlässigung der Modulationstransferfunktion eintretende Fehler würde bei einer Vergrößerung von 500 000 also etwa 13 % betragen. Da die in Abbildung 7.6 gezeigten Bilder nur Raumfrequenzen bis maximal 5.5 nm^{-1} enthalten, wäre der zu erwartende maximale Fehler bei Vernachlässigung der Modulationstransferfunktion nur noch 8 %. Eine Korrektur der Bildintensität wurde dennoch vorgenommen. Eine mögliche Abweichung der „wahren“ charakteristischen Raumfrequenz g_0 vom übernommenen Wert 40 mm^{-1} um bis zu 25 % hätte wegen der Geringfügigkeit der Korrektur kaum meßbare Auswirkungen.

Gemäß der im Theorieabschnitt beschriebenen Digitalisierungsmethode wird ein Ausschnitt des ursprünglich aufgenommenen Bildes aus 512×512 Bildpunkten so ausgewählt, daß darin eine ganze Zahl von Einheitszellen zu liegen kommt (Siehe Abb. 7.3). Die anfangs grob bestimmten Eckkoordinaten des ausgewählten Bildbereichs werden danach mittels Kreuzkorrelationsroutinen so optimiert, daß die wechselseitige Korrelation der im Bildbereich liegenden Einheitszellen maximal ist. Verzerrungen, die aufgrund der verschiedenen Ausdehnung der CCD-Anordnung in x- und y-Richtung, die Winkel zwischen orthogonalen Achsen, die nicht parallel zur Anordnung liegen, verfälschen, werden in diesem Bearbeitungsschritt automatisch eliminiert.

In einer zweiten Digitalisierungsstufe wird der so optimierte Bildbereich erneut auf einem Feld aus 256×256 Bildpunkten gespeichert. Die Intensitätswerte an den Bildpunkten des neuen Feldes werden aus denen des ursprünglichen mittels bilinearer Interpolation ermittelt. Die für diesen Bearbeitungsschritt gewählten Routinen gestatten auch den Einsatz von Interpolationen höherer als linearer Ordnung. Es zeigt sich jedoch, daß der Einsatz aufwendigerer Interpolationsmethoden, der bei „glatten“ Funktionen normalerweise einen Gewinn an Genauigkeit ermöglicht, keine Verbesserung bringt. Dies liegt daran, daß bei der zweiten Digitalisierung die Anzahl der Abtastpunkte pro Fläche größer als im ursprünglich aufgenommenen Bild ist, falls ein Bildausschnitt gewählt wird, der kleiner als ein Viertel des ursprünglichen Bildes ist. Es entsteht in solch einem Fall ein „Mangel“ an Stützpunkten, bei dem auch eine höhere Ordnung der Interpolation keinen Informationsgewinn ermöglicht.

Nach der Fouriertransformation des ausgewählten Bildbereichs wird der Ursprung der Fouriertransformierten im Falle eines zentrosymmetrischen Motivs automatisch um den anfänglich gewählten Nullpunkt innerhalb eines vorgegebenen Bereichs variiert, um den minimalen Wert des gewichteten Phasenresiduums (Gl. (7.19)) zu ermitteln. In einem zweiten Schritt wird bei vierzähliger Symmetrie des Bildmotivs der so gewonnene Ursprung nochmals verfeinert, indem der Bildkontrast minimiert wird, der von Fourierkoeffizienten stammt, die nicht die in Gleichung (7.22) aufgestellte Bedingung erfüllen. Nach diesem Schritt kann bei Bedarf die im Theorieteil beschriebene Methode zur Symmetriekorrektur des Bildes angewandt werden.

Die Qualität des zweiten Digitalisierungsschrittes und der mögliche Informationsverlust, der entsteht, wenn eine Filtermaske angewandt wird, die nur einzelne Fourierkoeffizienten „durchläßt“, kann auf zwei verschiedene Arten überprüft werden. Diese Möglichkeiten werden anhand der Abbildungen 7.4 und 7.5 verdeutlicht.

Der für den zweiten Digitalisierungsschritt gewählte Bildausschnitt wurde im Falle des in Abbildung 7.4 (a) gezeigten Bildes richtig gewählt und entspricht dem in Abbildung 7.3 gezeigten Rahmen. Das Bild beinhaltet exakt 3×3 Ni_4Mo Einheitszellen. Im Gegensatz dazu ist die Bildbegrenzung in Abbildung 7.4 (b) weder in ihrer Abmessung noch in ihrer Ausrichtung mit dem darin enthaltenen Bildmotiv kommensurabel. In der Fouriertransformierten des in Abbildung 7.4 (b) gezeigten Bildes kann deshalb ein deutliches „Verschmieren“ der Fourierkoeffizienten beobachtet werden. Das Ausmaß dieses unerwünschten Effekts ist jedoch anhand der Fouriertransformierten sehr schwer abzuschätzen.

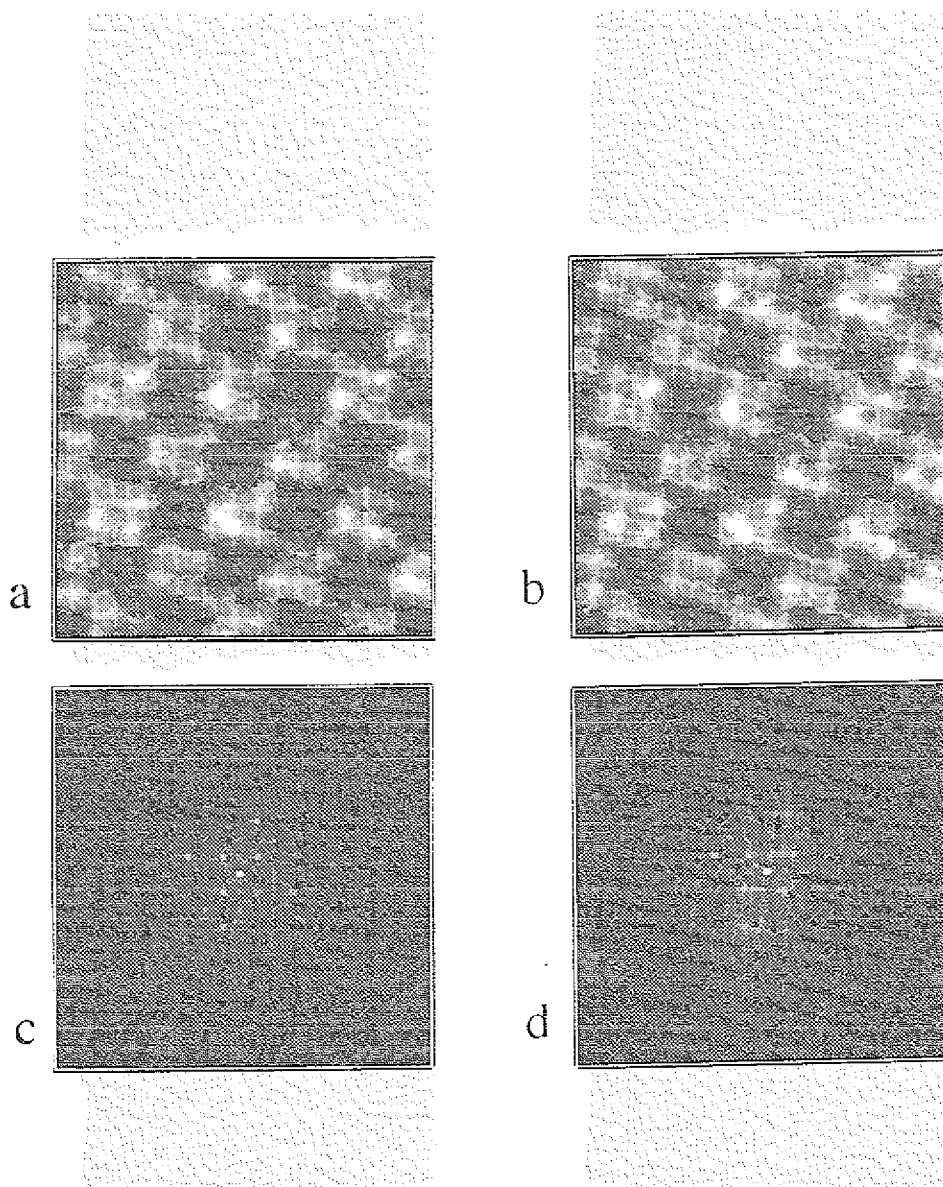


Abbildung 7.4: Oben: 3×3 Ni_4Mo Einheitszellen nach der zweiten Digitalisierungsstufe. Die gezeigten Bilder bestehen aus 256×256 Bildpunkten. Bild (a) entspricht dem in Abbildung 7.3 dargestellten Bildausschnitt, der genau 9 Einheitszellen beinhaltet. In Bild (b) sind weder die Länge noch die Ausrichtung des für die zweite Digitalisierungsstufe gewählten Rahmens mit der abgebildeten Struktur kommensurabel. Zudem ist das Bild leicht verzerrt. Unten: In den Teilen (c) und (d) der Abbildung ist der Betrag der zugehörigen Fouriertransformaten in logarithmischer Skala dargestellt. Während in der Fouriertransformaten (c) deutlich einzelne Pixel hoher Intensität sichtbar sind, erkennt man in (d) ein „Verschmieren“ der Intensitätsmaxima.

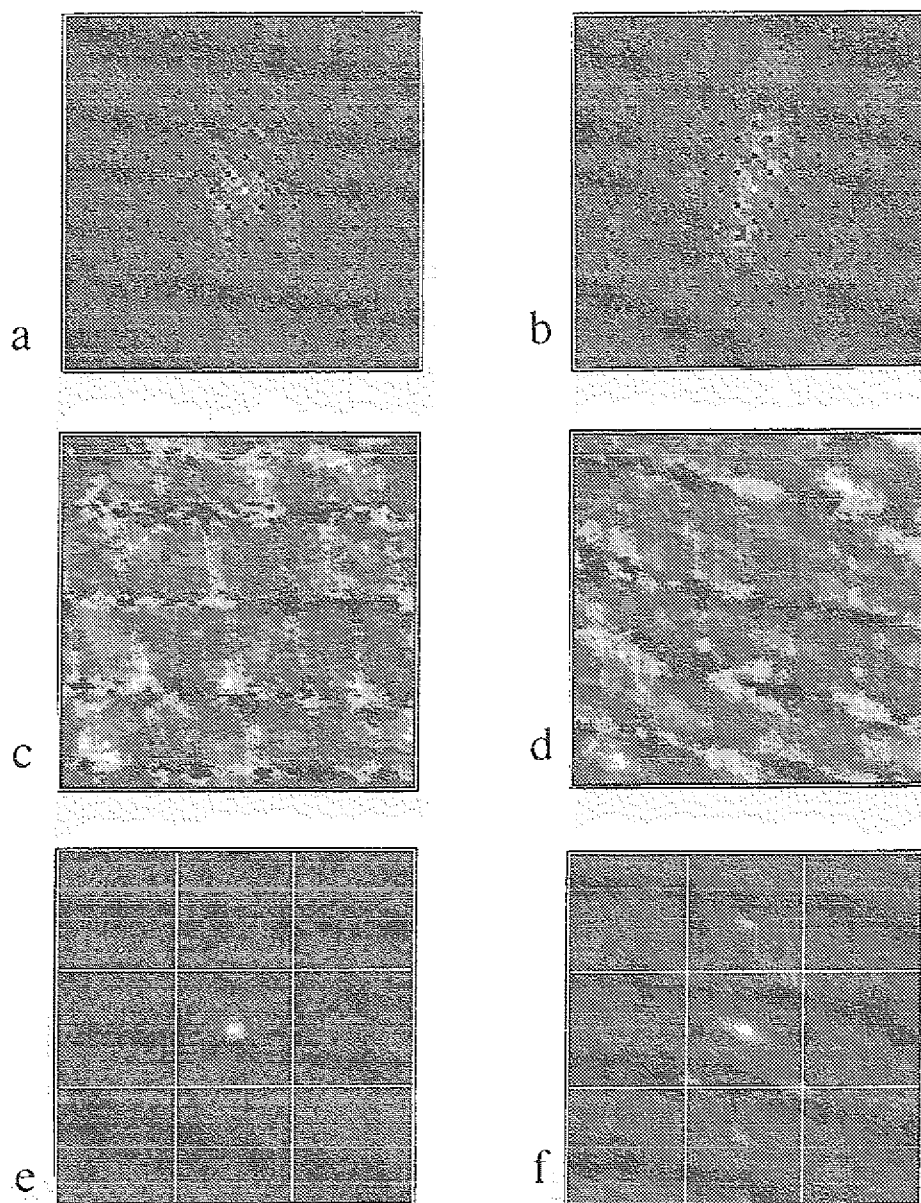


Abbildung 7.5: Linke und rechte Seite der Abbildung gehören zu den in Abbildung 7.4 (a) und (b) dargestellten Bildern. (a,b) : Ein Anti-Bragg Filter, bei dem die „abgedeckten“ Bereiche den Durchmesser eines einzelnen Pixels haben, wird auf die Fouriertransformierte angewandt. (c,d): Bilder, die aus der Rücktransformation der gefilterten Fouriertransformierten (a) und (b) hervorgehen und hauptsächlich Rauschen beinhalten. (e,f): Zentraler Ausschnitt aus der Autokorrelationsfunktion der Bilder (c) und (d). Die Linien markieren den Umriß der Ni_4Mo Einheitszelle. In (e) ist nur das zu erwartende zentrale Maximum der Autokorrelation zu sehen. Die zusätzlichen Maxima und Minima in (f) zeigen an, daß in Bild (d) neben dem Rauschen noch periodische Information verborgen ist. Die leichte Verzerrung des Ausgangsbildes in Abbildung 7.4 (b) ist deutlich an der elliptischen Form des zentralen Maximums der Autokorrelation (f) erkennbar.

Eine andere Möglichkeit, den durch die Anwendung des Bragg-Filters entstehenden Informationsverlust sichtbar zu machen, besteht in der Betrachtung der bei der Filterung „verworfenen“ Bilddetails. Ein sogenannter Anti-Bragg Filter, der zum Ausgangsfilter komplementär ist, wird auf die Fouriertransformierte der in Abbildung 7.4 gezeigten Bilder angewandt (Siehe Abb. 7.5 (a) und (b)). Alle Raumfrequenzen, die vorher als Rauschen interpretiert wurden, werden von diesem Komplementärfilter hindurchgelassen und in den Realraum zurücktransformiert. Das Ergebnis der Rücktransformation ist in den Abbildungen 7.5 (c) und (d) gezeigt. Eine nicht mit dem Bildmotiv kommensurable Wahl des Bildausschnittes „verrät“ sich in der hauptsächlich aus Rauschen bestehenden Rücktransformierten durch das zusätzliche Auftreten von periodischen Kontrastmodulationen. Sehr schwache Kontrastdetails, die im Rauschen oft schwierig zu entdecken sind, können durch die Bildung der Autokorrelation (Siehe Abb. 7.5 (e) und (f)) der rücktransformierten Bilder sichtbar gemacht werden [92]. Die Autokorrelation eines Bildes, das aus unkorreliertem Rauschen besteht, weist ausschließlich ein zentrales Maximum auf. Ist jedoch periodische Information im Rauschen verborgen, so treten in der Autokorrelation noch zusätzliche Maxima und Minima auf. Solche zusätzlichen Maxima und Minima sind in Abbildung 7.5 (f) deutlich erkennbar. In diesem Fall würde also Information durch die Anwendung eines Bragg-Filters verloren gehen. Mit Hilfe dieser Kontrollmethode läßt sich genau überprüfen, ob der mittels des Bragg-Filters extrahierte Satz von Fourierkoeffizienten ausreichend genau das experimentelle Bild repräsentiert.

Der so gewonnene Satz von experimentellen Fourierkoeffizienten wird an das Hauptprogramm des Pakets übergeben, das bei systematischer Variation von Defokus und Probendicke eine große Anzahl von Bildern berechnet und mit den experimentellen Eingabebildern vergleicht. Durch die Berechnung von typischerweise 10^4 Bildern ist ein quasikontinuierliches Abtasten des gesamten relevanten Bereichs der Defokus-Dicken-Ebene möglich. Die Fourierdarstellung jedes einzelnen simulierten Bildes wird direkt nach seiner Berechnung mit den Fourierkoeffizienten des experimentellen Eingabebildes verglichen. Die aus dem Vergleich resultierenden Kreuzkovarianzkoeffizienten werden über der Defokus-Dicken-Ebene aufgetragen (Siehe Abb. 7.8).

Die hierfür erforderliche Rechenzeit hängt von der Anzahl der Strahlen, die innerhalb der Objektivblende liegen, sowie von der Anzahl der zu vergleichenden simulierten und experimentellen Bilder ab. Bei den in diesem Kapitel gezeigten Bildern der Legierung Ni_4Mo beträgt die Anzahl der Strahlen 29, der Defokus wurde in Schritten von 1 nm und die Probendicke in Schritten von 0.5 nm variiert. Jeweils 100 Defokus- und 100 Dickenwerte wurden in die Simulation einbezogen. Die Rechenzeit für die Simulation von 10^4 Bildern sowie für den anschließenden Vergleich dieser Bilder mit vier experimentellen Eingabebildern beträgt auf einer DEC VAX station 3100 etwa 5 Minuten. Auf einer bei weitem schnelleren DEC 3500 ALPHA Rechenanlage reduziert sich die Rechenzeit auf etwa 10 Sekunden.

7.4 Defokus- und Probendickenbestimmung

Ein idealer Test für die quantitative Bilderkennungsmethode besteht in der Lokalisierung einer Defokus-Serie in der Defokus-Dicken-Ebene. Bei erfolgreicher Bilderkennung sollten die gefundenen Werte für die Probendicke für alle Bilder der Serie übereinstimmen.

In der linken Spalte von Abbildung 7.6 ist eine Defokus-Serie aus vier experimentellen Bildern der Legierung Ni_4Mo nach der zweiten Digitalisierungsstufe dargestellt. Jedes dieser Bilder besteht aus 256×256 Bildpunkten und zeigt 4×4 Einheitszellen in der $[001]$ Projektionsrichtung. Der Kontrast der Bilder 1–3 wird von der Periodizität der Molybdän-Überstruktur dominiert. In Bild 4 kann man Kontrastmodulationen erkennen, die die Periodizität der kfz Matrix besitzen. In den Bildern 3 und 4 treten starke Abweichungen von der vierzähligen Symmetrie der D1_a Struktur auf. Zusätzlich ist der Kontrast von Bild 2 ungleichmäßig, wobei an manchen Stellen die zwischen größeren hellen Punkten liegenden feineren Kontrastmaxima (engl. *half spacing contrast*) verschwinden.

Für die Bildsimulation wurde als Koordinatenursprung eine Molybdänposition ausgewählt. Der Bildvergleich mittels des Kreuzkovarianzkoeffizienten erfordert, daß das zu vergleichende experimentelle Bild denselben Ursprung hat, d.h. daß die den Bildern zugrundeliegende Struktur zur Deckung gebracht werden muß. Hierbei entsteht ein nicht-triviales Problem, da von vornherein nicht klar ist, welcher Bildpunkt im experimentellen Bild einer Molybdänposition im simulierten Bild entspricht. Zum Beispiel könnten sich in Bild 1 aus Abbildung 7.6 die Molybdänpositionen an den Intensitätsmaxima oder an den dazwischenliegenden Minima befinden. Die gleiche Problematik entsteht auch bei den restlichen Bildern, es gibt immer jeweils zwei Möglichkeiten der Positionierung. Der Grund hierfür liegt darin, daß die D1_a Struktur zwei Zentren vierzähliger Symmetrie besitzt, wovon eine von einem Molybdänatom besetzt ist, die andere jedoch unbesetzt ist. Die einzige Möglichkeit, das daraus resultierende Problem zu lösen, besteht darin, beide Möglichkeiten der Ursprungswahl in den Bildvergleich einzubeziehen. Die „falsche“ Variante läßt sich bei einer hinreichend komplexen Struktur, wie Ni_4Mo sie darstellt, dadurch ausschließen, daß sich bei der Durchsuchung der Defokus-Dicken-Ebene kein dazu passendes simuliertes Bild findet. Dies ist bei jedem der vier in Abbildung 7.6 gezeigten Bildern der Fall, sodaß jede Mehrdeutigkeit ausgeschlossen werden kann.

In der mittleren Spalte von Abbildung 7.6 sind Bilder der Defokus-Serie nach Anwendung des Bragg-Filters zu sehen. Die zusätzlich der Symmetriekorrektur unterzogenen Bilder der Serie sind in der rechten Spalte dieser Abbildung gezeigt. Im vierten Bild der rechten Spalte sind aufgrund der nur partiellen Symmetriekorrektur noch leichte Asymmetrien im Bildkontrast zu sehen. Aufgrund von Abweichungen der Phasenwinkel von mehr als 90° sind die $\{220\}$ Koeffizienten dieses Bildes nicht korrigiert worden (Siehe Gl. (7.22)). Die partiell symmetriekorrigierten Bilder 1–4 sind nochmals in der linken Spalte von Abbildung 7.7 zusammen mit den vom Bilderkennungsprogramm ermittelten simulierten Bildern gezeigt, die dazu die beste Übereinstimmung zeigen.

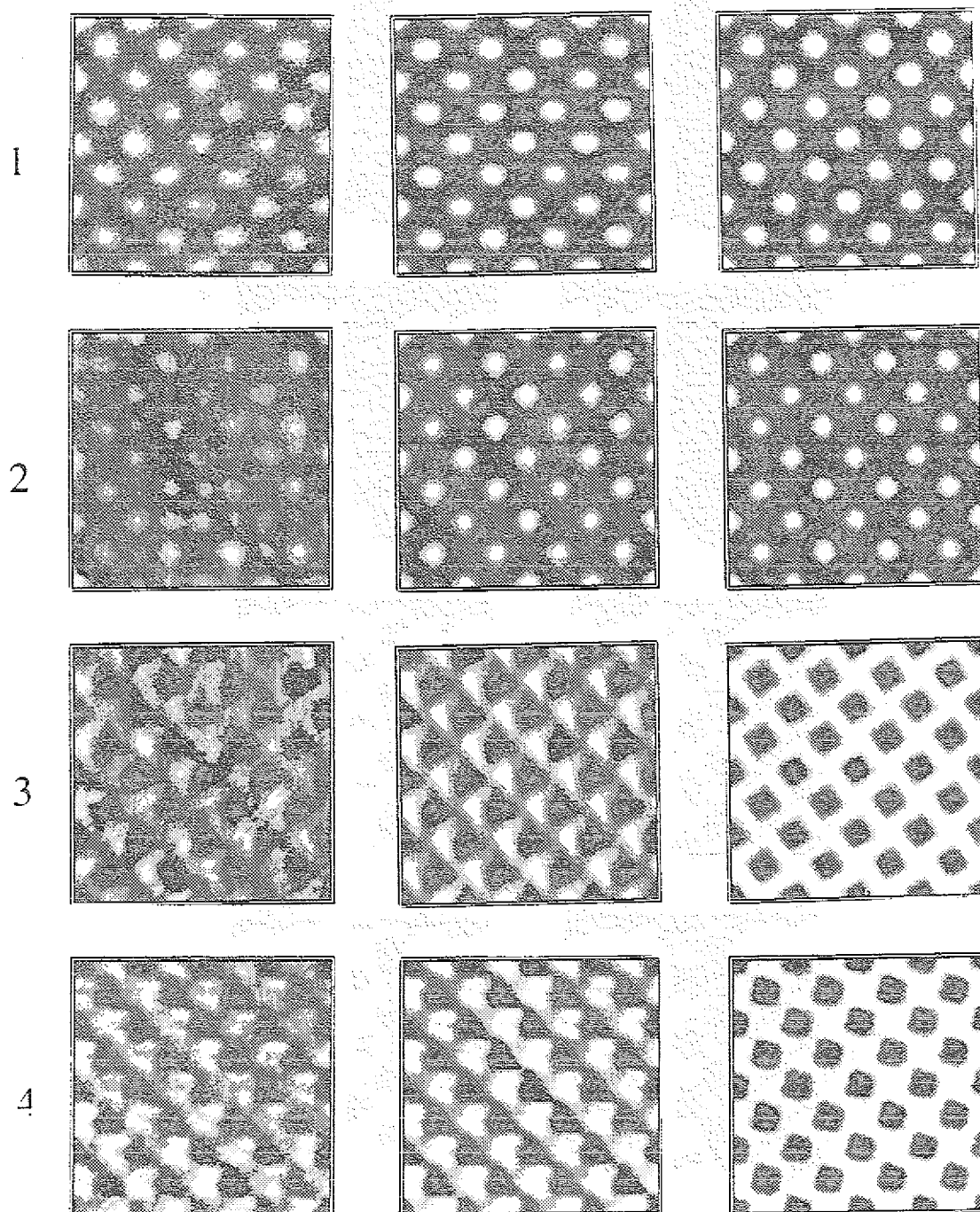


Abbildung 7.6: Entlang der [001] Zonenachse aufgenommene Defokus-Serie der Legierung Ni_4Mo . Jedes Bild besteht aus 256×256 Bildpunkten und enthält 4×4 Einheitszellen. Die Ecken der dargestellten Bilder entsprechen Molybdänpositionen. Linke Spalte: Bilder 1–4 nach der zweiten Digitalisierungsstufe. Mittlere Spalte: Bilder 1–4 nach Anwendung des Bragg Filters. Rechte Spalte: Partiell symmetriekorrigierte Bilder 1–4.

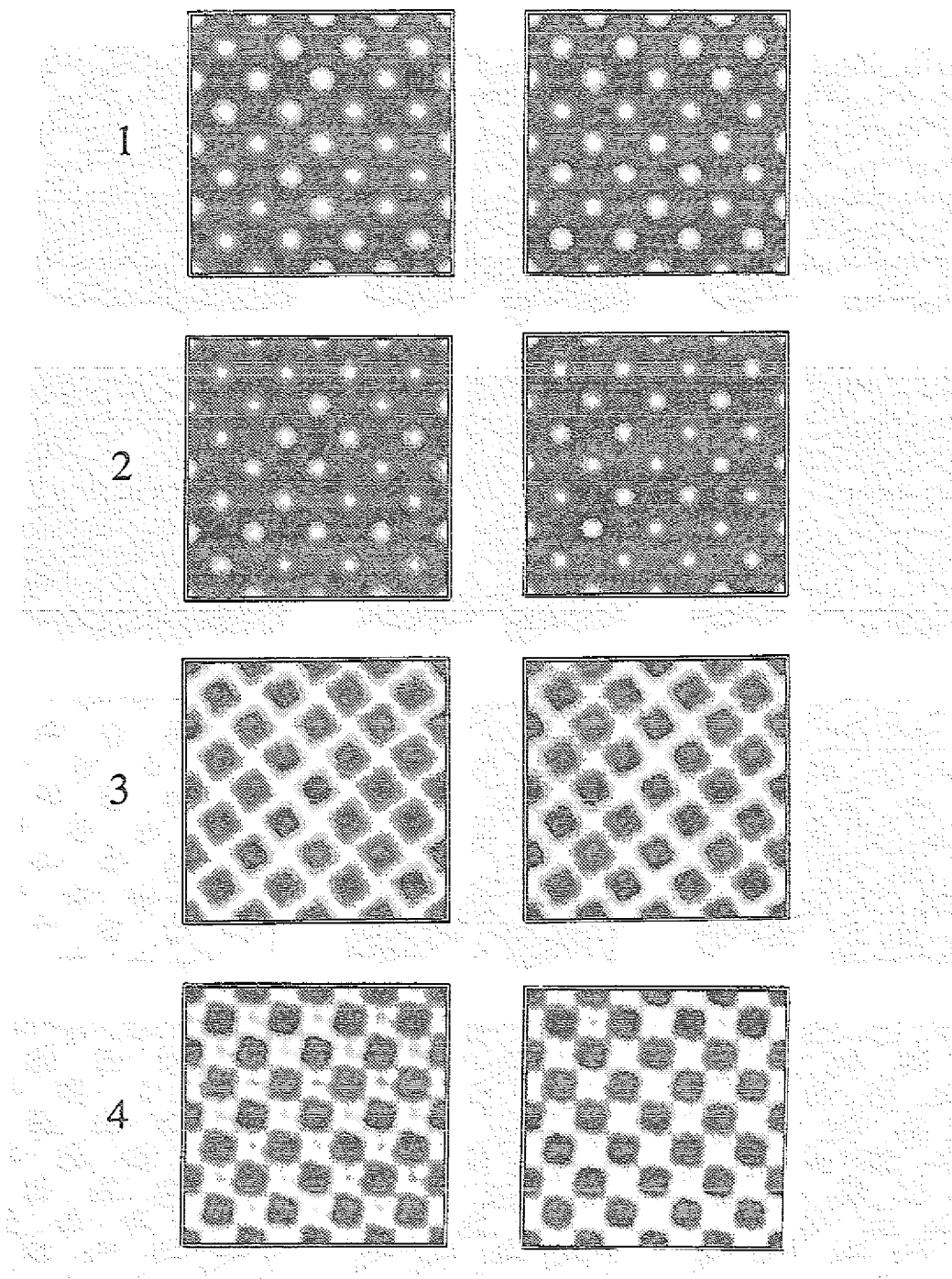


Abbildung 7.7: Linke Spalte: Bilder 1–4 der Ni_4Mo Defokus-Serie nach Anwendung des Bragg-Filters sowie des Verfahrens zur Symmetriekorrektur (gleich wie rechte Spalte in Abbildung 7.6). Rechte Spalte: Die am besten übereinstimmenden simulierten Bilder, die vom Mustererkennungsprogramm gefunden wurden. Die zugehörigen Werte für Defokus und Probendicke sind in Tabelle 7.1 aufgelistet.

Bild Nr.		1	2	3	4
Min. Phasenresiduum /[Grad]		8.4	13.6	26.5	26.0
Max. Kovarianz /[-]	a	0.979	0.954	0.904	0.890
	b	0.995	0.999	0.997	0.990
Defokus /[nm]	a	26	15	-28	-25
	b	25	14	-48	-57
Probendicke /[nm]	a	14	13	14	14
	b	14	13.5	14	14.5
a: Nur Bragg Filterung, b: Filterung + Symmetriekorrektur					

Tabelle 7.1: Werte minimalen Phasenresiduums für die in der linken Spalte von Abbildung 7.6 gezeigten vier experimentellen Bilder sowie Kovarianzkoeffizient, Defokus und Probendicke der damit am besten übereinstimmenden simulierten Bilder.

Die Werte minimalen Phasenresiduums der vier Bilder aus der Defokus- Serie sind in Tabelle 7.1 aufgelistet. Die starke Abweichung der Bilder 3 und 4 von der zweizähligen Symmetrie spiegelt sich in den relativ hohen Phasenresiduen um 25° wider. In Tabelle 7.1 sind auch die Werte der Kreuzkovarianz der am besten übereinstimmenden simulierten Bilder mit den zugehörigen Werten für Defokus und Probendicke dargestellt. Der erste Satz von Werten (a) bezieht sich auf experimentelle Eingabebilder, die nur gefiltert sind jedoch nicht der Symmetriekorrektur unterzogen wurden (Abb. 7.6 mitte). Der zweite Satz von Werten (b) bezieht sich auf die zusätzlich symmetriekorrigierten Varianten der Eingabebilder (Abb. 7.6 rechts). Während die Symmetriekorrektur keinen wesentlichen Einfluß auf die bei den Bildern 1 und 2 erzielten Ergebnisse hat, ändern sich die bei den Bildern 3 und 4 gefundenen Defokuswerte beträchtlich, je nachdem ob die Symmetriekorrektur durchgeführt oder weggelassen wurde. Bei Anwendung der Symmetriekorrektur verschieben sich die gefundenen Defokuswerte in Richtung Unterfokus. Dieser Sachverhalt wird in der Diskussion näher erläutert.

Wie an früherer Stelle bereits vermerkt, ist es bei den hier gezeigten Bildern eindeutig möglich, von der Lage des Koordinatenursprungs der simulierten Bilder auf den entsprechenden Ursprung der experimentellen Bilder zu schließen, da es bei keinem der experimentellen Bilder vorkommt, daß *beide* alternativ möglichen Ursprungsvarianten hohe Werte der Kovarianz zu simulierten Bildern liefern. Somit ist es eindeutig möglich, die Atompositionen in den experimentellen Bildern zu bestimmen. Die deutlich sichtbaren Intensitätsmaxima in den Bildern 1 und 2 der Serie liegen an projizierten Molybdänpositionen. Im Gegensatz dazu liegen in den Bildern 3 und 4 die Molybdänpositionen an Intensitätsminima.

Die Werte für die Probendicke der am besten zu den experimentellen Bildern korrelierten Simulationsbilder variieren unabhängig von der Anwendung der Symmetriekorrektur nur um 1.5 nm. Diese Schwankung entspricht der „Höhe“ von 4 Einheitszellen.

Es ist jedoch unter Umständen nicht ausreichend, Defokus und Probendicke von einzelnen simulierten Bildern abzuleiten, die maximale Kreuzkovarianz zu experimentellen Bildern aufweisen. Ähnliche oder beinahe gleiche Bilder, die nur einen leicht geringeren Wert der Kreuzkovarianz ergeben, können bei völlig anderen Kombinationen von Defokus und Probendicke auftreten. Um die Eindeutigkeit der gefundenen Werte zu überprüfen, wird eine Kreuzkovarianz-Karte erstellt. In Abbildung 7.8 sind in der Defokus-Dicken-Ebene Linien gleicher Kreuzkovarianz zwischen den symmetriekorrigierten experimentellen Bildern und simulierten Bildern auf den Niveaus $c = 0.98$ und $c = 0.99$ dargestellt.

In der Kreuzkovarianz-Karte treten scharfe Maxima bei den zwei Unterfokusbildern 3 und 4 auf. Da diese Bilder aus der Interferenz einer großen Zahl von Strahlen entstehen, die unterschiedliche dynamische Extinktionslängen sowie unterschiedliche Wiederholungsperioden der Aberrationsfunktion besitzen (Gl. 5.27 sowie [93]), ist ihr Auftreten in der Defokus-Dicken-Ebene sehr stark lokalisiert. Für Bild 3 sind zwei Defokuswerte, nämlich -30 nm und -48 nm, möglich, die Probendicke von 14 nm ist jedoch eindeutig bestimmbar.

Im Gegensatz dazu sind die Bereiche hoher Kreuzkovarianz bei den Bildern 1 und 2 viel großflächiger in der Defokus-Dicken-Ebene verteilt. Diese Bereiche liegen fast ausschließlich im Überfokus. Bei solchen Fokussierungsbedingungen werden aufgrund der partiellen räumlichen Kohärenz des einfallenden Elektronenstrahls nur wenige Strahlen mit geringer Auflösung vom Mikroskop übertragen (Siehe Kapitel 5). Diese Strahlen besitzen eine äußerst lange Wiederholungsperiode $\Delta Z = 2/(g^2\lambda)$ in der Aberrationsfunktion. Zum Beispiel wiederholt sich der lineare Bildkontrast der $\{110\}$ Strahlen, die den Kontrast der Bilder 1 und 2 dominieren, erst nach einer Änderung des Defokus von 199 nm. Dieses systematische Verhalten kann deutlicher in der Kovarianz-Karte, die im nächsten Abschnitt gezeigt wird, beobachtet werden.

7.5 Kontrastübersicht

Die in den vorangehenden Abschnitten beschriebene Methode zur quantitativen Bilderkennung bietet die Möglichkeit, das Auftreten charakteristischer Bildtypen in der Defokus-Dicken-Ebene systematisch zu erfassen. Statt experimenteller Bilder können auch simulierte Bilder, definiert durch entsprechende Defokus- und Dickenwerte, als Eingabe zur Mustererkennung dienen.

Sechs intuitiv leicht interpretierbare simulierte Bilder von Ni_4Mo , das entlang der $[001]$ Zonenachse orientiert ist, werden als Referenzbilder benutzt. In Bild 1 der Abbildung 7.9 erscheinen projizierte Mo-Positionen als helle Punkte, während in Bild 3 sich die hellen Punkte an sogenannten Tunnelpositionen zwischen Mo-Atomen befinden. Der Kontrast in Bild 2 ist eine „Mischung“ zwischen den beiden erstgenannten Bildern, es tritt ein sogenannter „halfspacing“ Kontrast der Überstruktur auf. Bild 4 ist dadurch gekennzeichnet, daß an den Mo-Positionen dunkler Kontrast auftritt, während Ni-Positionen hell erscheinen. In Bild 5 sind die Gitterpunkte des projizierten kfz Matrixgitters hell, Ni- und Mo-Atome sind hier nicht unterscheidbar. Bild 6 entspricht einer Umkehr des Kontrasts im Bezug auf das vorangehende

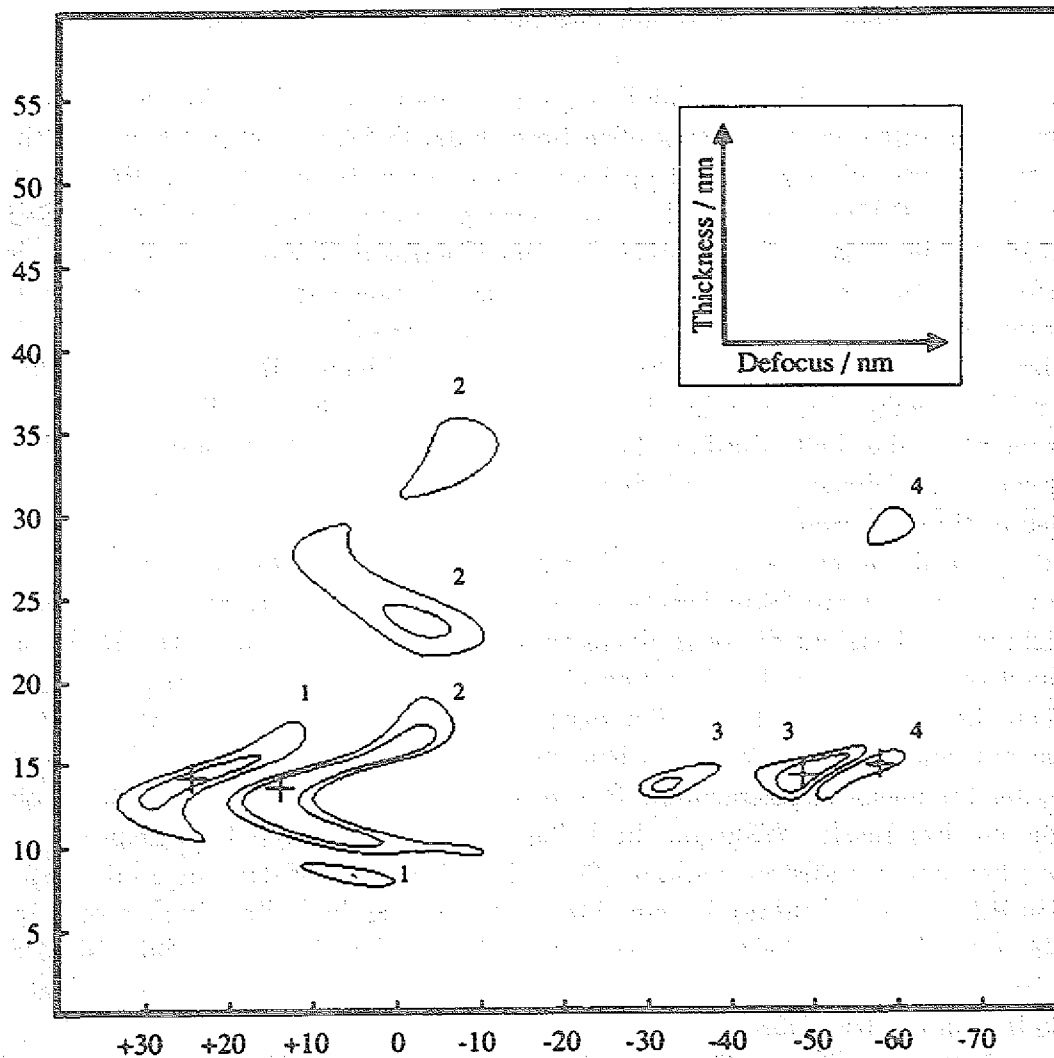


Abbildung 7.8: Kreuzkovarianz-Karte. Linien gleicher Kreuzkovarianz zwischen simulierten Bildern und den partiell symmetriekorrigierten Bildern 1– 4 der Ni_4Mo Defokus-Serie sind in der Defokus-Dicken-Ebene aufgetragen. Die Kreuzkovarianz beträgt an den äußeren Linien 0.98. Innere Linien – falls vorhanden – entsprechen einem Wert von 0.99. Die Positionen maximaler Kreuzkovarianz sind durch Kreuze gekennzeichnet. Die Zahlen beziehen sich auf die in Abb. 7.7 gezeigten Bilder.

Bild. Da der Kontrast dieser sechs ausgewählten Bilder eng mit der projizierten Struktur zusammenhängt, werden diese im folgenden als Bildtypen bezeichnet. In Abbildung 7.9 sind alle Gebiete der Defokus-Dicken-Ebene gekennzeichnet, wo simulierte Bilder eine Kreuzkovarianz zu einem der sechs Referenzbilder haben, die größer als 0.95 ist. Der Wert 0.95 erweist sich als gutes Maß um den subjektiven Eindruck zu beschreiben, ob zwei Bilder ungefähr gleich oder deutlich unterschiedlich sind.

In Abbildung 7.9 sieht man, daß Bildtyp 1, bei dem nur an Mo- Positionen Intensitätsmaxima auftreten, in einem großen Bereich der Defokus-Dicken-Ebene auftritt. In Richtung Unterfokus geht Bildtyp 1 kontinuierlich in Bildtyp 2 über. Beide Bildtypen, die den Bildern 1 und 2 der zuvor gezeigten experimentellen Defokus-Serie entsprechen, „bedecken“ einen großen Teil der Überfokus-Region sowie den Teil der Unterfokus-Region, der sich bis zu einer Defokussierung von -20 nm erstreckt. Im Gegensatz dazu ist Bildtyp 3, der durch Intensitätsmaxima an den Tunnelpositionen der Überstruktur gekennzeichnet ist, auf einen kleinen Bereich der Defokus-Dicken-Ebene nahe einer Defokussierung von -75 nm beschränkt. Es ist daher sehr wahrscheinlich, daß helle Punkte, die in einer experimentellen Abbildung ein Quadratmuster mit Überstrukturperiodizität bilden, an Mo- Positionen und nicht an Tunnelpositionen liegen.

Bildtyp 4, der es ermöglicht, zwischen Ni- und Mo- Positionen zu unterscheiden, tritt nur bei einer Unterfokussierung von etwa -60 nm auf. Im Hinblick auf die Proben Dicke ist die Existenz dieses Bildtyps auf drei „Inseln“ beschränkt. Der Dickenunterschied zwischen diesen Inseln entspricht ungefähr der Extinktionslänge der $\{200\}$ Strahlen der kfz Matrix. Bild 4 der experimentellen Defokus-Serie (Abb. 7.6) erscheint auf den ersten Blick diesem Bildtyp sehr ähnlich. Der Wert für die Proben Dicke, der für dieses experimentelle Bild ermittelt wurde, fällt jedoch nicht in eine von diesen drei Inseln. Während die hellen Punkte in Bildtyp 4 Nickelpositionen entsprechen und äquidistant entlang $\{200\}$ Ebenen der kfz Matrix angeordnet sind, kann in Bild 4 der Abbildung 7.6 eine klare Abweichung der hellen Punkte von einer äquidistanten linearen Positionierung beobachtet werden. In diesem Bild fallen die Maxima nicht mit Atompositionen zusammen und eine intuitive Bildinterpretation würde hier in die Irre führen.

Unterhalb einer Proben Dicke von 5 nm ist deutlich die Wiederholungsperiode der Aberrationsfunktion bezüglich der $\{200\}$ Strahlen der kfz Matrix zu sehen. In Abbildung 7.9 sieht man deutlich, wie sich die Bildtypen 5 und 6 mit einem Defokusabstand von 20 nm wiederholen. Dies entspricht fast genau der halben Wiederholungsperiode $\Delta Z = 2/(g^2 \lambda)$ der $\{200\}$ Strahlen der kfz Matrix, die 39.8 nm beträgt. Die Tatsache, daß die Matrix nur bei Proben Dicken unterhalb 5 nm beobachtet werden kann, beruht auf der großen Extinktionslänge der Überstrukturreflexe. Bei geringen Proben Dicken haben die Überstrukturreflexe noch nicht genügend Intensität gewonnen, um in den Bildern beobachtet werden zu können (Siehe Abb. 4.5).

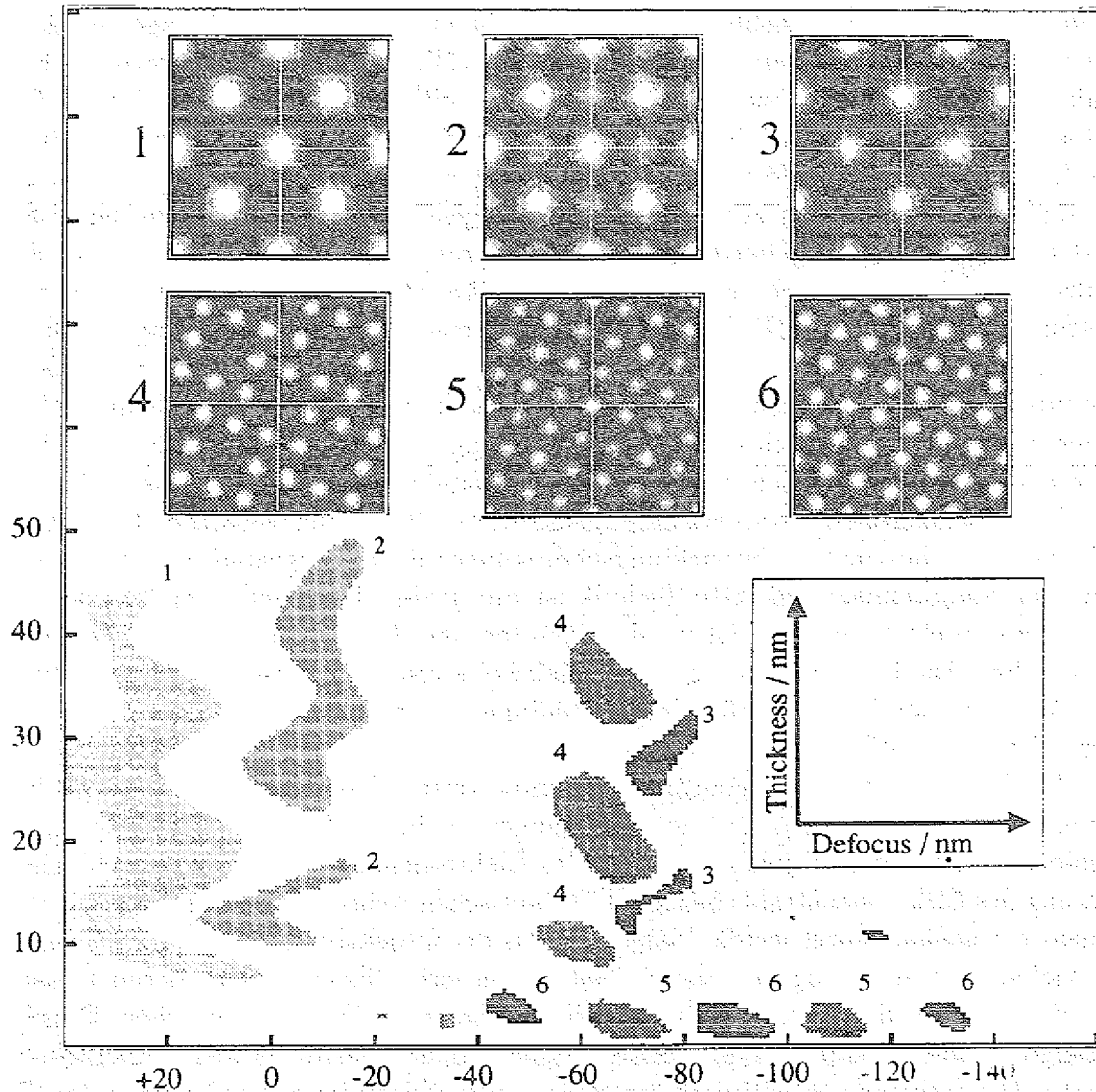


Abbildung 7.9: Kreuzkovarianz-Karte. Die sechs dargestellten simulierten Bilder bestehen aus jeweils 2×2 Einheitszellen und dienen als Referenzbilder zur Mustererkennung. Gebiete der Defokus-Dicken-Ebene, wo simulierte Bilder eine Kreuzkovarianz größer als 0.95 zu den Referenzbildern haben, sind durch verschiedene Graustufen gekennzeichnet. Die Zahlen in der Karte beziehen sich auf die oben dargestellten Referenzbilder.

7.6 Diskussion

Die quantitative Bilderkennungsmethode, die in den vorangehenden Abschnitten vorgestellt wurde, umfaßt zwei Stufen. Dies sind zum einen die Extraktion von Fourierkoeffizienten aus experimentellen Aufnahmen, zum anderen der darauf folgende Vergleich dieser Koeffizienten mit Fourierkoeffizienten simulierter Abbildungen. Die in der zweiten Stufe erreichbare Genauigkeit ist nur durch die Rechengenauigkeit des verwendeten Computers begrenzt und somit völlig problemlos. Die korrekte Extraktion von Fourierkoeffizienten ist eine weitaus diffizilere Aufgabe, die jedoch für Bilder, die ein gewisses Minimum an Qualität besitzen, mit den hier angewandten Methoden quantitativ gelöst und auch überprüft werden kann.

Im Hinblick auf einen quantitativen Bildvergleich ist die Verwendung einer direkt am Mikroskop installierten CCD- (*charge coupled device*) Kamera [68, 6, 69] wünschenswert, die jedoch zum Zeitpunkt der Anfertigung dieser Arbeit nicht zur Verfügung stand. Eine CCD- Kamera ermöglicht eine lineare Aufzeichnung der Elektronenintensität, zudem ist der erfassbare Dynamikbereich der Intensität um Größenordnungen besser als der einer photographischen Emulsion. Eine CCD- Kamera hat einen dynamischen Bereich in der Größenordnung von 10^4 Graustufen oder 14 bit, während Photoplatten nicht mehr als 256 Graustufen bzw. 8 bit auflösen können. Der Einsatz einer CCD- Kamera erspart die Anwendung einer Software, die maßgeschneidert für bestimmte Entwicklungsbedingungen des photographischen Films ist. Die hier *notgedrungen* benutzte Technik ist mit großen Unsicherheiten verbunden, da Effekte nicht berücksichtigt werden konnten, die die Ergebnisse möglicherweise verfälschen: Die Filmschwärzung ist beispielsweise von der Konzentration und somit der Alterung der Entwicklerflüssigkeit abhängig, ebenso spielt die Lagerungsdauer der Filme eine Rolle.

Bei der Extraktion der Fourierkoeffizienten treten Probleme bei solchen Abbildungen auf, in denen Symmetrieabweichungen zu beobachten sind, die von einer Fehlorientierung des Objekts bezüglich des Elektronenstrahls, von einer Fehlorientierung des Elektronenstrahls bezüglich der optischen Achse des Mikroskops oder von einem mangelhaft korrigierten Astigmatismus der Objektivlinse herrühren können. In vielen Hochauflösungsaufnahmen treten Fourierkoeffizienten auf, deren Phasen beträchtlich von den durch die Symmetrie vorgegebenen Werten abweichen. Dies ist häufig bei Raumfrequenzen der Fall, die nahe an oder jenseits der Scherzerauflösung liegen. Die steigende Sensitivität der Phasen mit steigender Raumfrequenz der abgelenkten Strahlen ist leicht zu erklären. Im Falle einer Fehlorientierung des Elektronenstrahls bezüglich der Objektivlinse steigt der dadurch induzierte phasenschiebende Beitrag zur Aberrationsfunktion mit der dritten Potenz der entsprechenden Raumfrequenz an [94]. Der Wert des Defokus, der im Falle von Astigmatismus in zwei orthogonalen Raumrichtungen unterschiedlich ist, trägt mit dem Quadrat der Raumfrequenz zur Aberrationsfunktion bei [62]. Ebenso ist es intuitiv einleuchtend, daß eine Fehlorientierung des Objekts hohe Raumfrequenzen stärker beeinflußt als niedrige.

Es ist möglich, die experimentellen Parameter, die zur Verschlechterung der Bilder geführt haben, im nachhinein zu bestimmen. Im schlimmsten Fall – wenn alle

der drei oben genannten Effekte auftreten – müssen neben Defokus und Proben-
dicke noch sechs weitere Parameter bestimmt werden. Im Falle von Strahl- oder Ob-
jektfehlorientierung, müssen jeweils die Richtung sowie die Stärke der Verkipfung
ermittelt werden, was die Bestimmung von insgesamt vier Parametern erfordert.
Astigmatismus wird durch die Raumrichtung des durchschnittlichen Defokus, sowie
durch die maximale Abweichung von diesem Durchschnittswert definiert [62], was
nochmals die Bestimmung von zwei Parametern erforderlich macht. Eine genaue
Bestimmung dieser sechs Parameter – ein Beispiel dafür ist in [95] zu finden – ist
äußerst aufwendig und wird dadurch kompliziert, daß die drei genannten Phänomene
nicht unabhängig voneinander behandelt werden können.

Die in dieser Arbeit angewandte Symmetriekorrekturemethode ersetzt nicht die
Suche nach den Parametern, die die unerwünschten Bildverzerrungen hervorrufen.
Die Methode ist nicht in der Lage, die „richtigen“ Bilder, die man ohne die ex-
perimentellen Fehler gemacht hätte, exakt zu rekonstruieren, obwohl dies offen-
bar näherungsweise doch möglich ist (Siehe Abb. 7.7). Ein Grund, warum diese
Methode zwar die Symmetrie jedoch nicht unbedingt das Bild korrigiert, liegt in
der nichtlinearen Charakteristik der Kontrastentstehung. Sieht man einmal von den
Dämpfungstermen aufgrund der partiellen Kohärenz ab, so setzt sich jeder Fourier-
koeffizient des Bildes aus einer großen Zahl von Interferenztermen zusammen. Geht
man von N Strahlen innerhalb der Objektivblende aus, so liegt diese Zahl in der
Größenordnung von N . Die oben genannten experimentellen Fehler wirken sich auf
jeden dieser Beiträge *unterschiedlich* aus, *bevor* sie in der komplexen Zahlenebene
zum Fourierkoeffizienten des Bildes addiert werden. Es ist jedoch im Sinne einer
Bildwiederherstellung nicht möglich, die Summe dieser komplexen Beiträge mittels
einer *gemeinsamen* Phasendrehung *nach* der Bildentstehung wieder zu kompensie-
ren. Ein weiterer Grund, warum die Korrekturmethode nur die Symmetrie, jedoch
nicht unbedingt das Bild wiederherstellt, läßt sich für den Fall der Fehlorientierung
des Objekts leicht angeben: Der dynamische Beugungsprozeß ändert sich bei Kip-
pung des Objekts aus der Zonenachse auf sehr komplexe Art und Weise. Dabei
sind *alle* Strahlen innerhalb des Objekts gekoppelt. Zur Korrektur können jedoch
nur Intensitätsbeiträge von solchen Strahlen herangezogen werden, die innerhalb der
Objektivblende liegen.

Dennoch spielt die Symmetriekorrektur bei der Mustererkennung eine entschei-
dende Rolle. Im Theorieabschnitt wurde anhand von Abbildung 7.1 anschaulich
erklärt, wie sich die Verschiebung von experimentellen Bildinhalten gegenüber dem
simulierten Idealbild bei der Mustererkennung mittels der Kreuzkovarianz auswirkt:
Simulierte Bilder, in denen die entsprechenden Bildinhalte erst gar nicht auftreten,
werden als besser korreliert zum „fehlerhaften“ experimentellen Bild eingestuft als
solche, bei denen die fraglichen Details sowohl vorhanden als auch richtig positio-
niert sind. Es ist einleuchtend, daß der Wert des Defokus, der ja aus der besten
Übereinstimmung des experimentellen Bildes mit einem simulierten Bild bestimmt
wird, fehlerhaft ist, wenn die beste Übereinstimmung beim „falschen“ simulierten
Bild festgestellt wird, das durch das Fehlen entscheidender Kontrastdetails gekenn-
zeichnet ist.

Da höher aufgelöste Bilddetails stärker zur Verschiebung neigen als solche mit niedriger Auflösung, besteht also die Tendenz, daß die Kreuzkovarianz simulierte Bilder „bevorzugt“, die geringere Auflösung besitzen als das „fehlerhafte“ experimentelle Eingabebild. Dies führt wiederum zu einer systematischen Verschiebung der gefundenen Defokuswerte in Richtung Überfokus. Aufgrund der Eigenschaften der Dämpfungenvelope, die den Einfluß der partiellen räumlichen Kohärenz des Elektronenstrahls beschreibt, besitzen Bilder, die im Überfokus oder im niedrigen Unterfokus aufgenommen wurden, eine geringere Auflösung als solche, die deutlich im Unterfokus aufgenommen wurden [53].

Das beschriebene Verhalten ist deutlich an den Kovarianzwerten in Tabelle 7.1 zu sehen. Während bei den Bildern 1 und 2 der in Abbildung 7.6 gezeigten Defokus-Serie die Symmetriekorrektur kaum einen Einfluß auf die ermittelten Defokuswerte hat, ändern sich bei den Bildern 3 und 4 nach der Symmetriekorrektur die ermittelten Defokuswerte drastisch in Richtung Unterfokus. Die starken Symmetrieabweichungen in den Bildern 3 und 4 spiegeln sich in den hohen Phasenresiduen um 25° wider. Ohne Symmetriekorrektur hat der Mustererkennungsalgorithmus für die beiden experimentellen Bilder nahezu identische Werte für Defokus und Dicke ermittelt, was offensichtlich falsch ist, da die beiden Eingabebilder deutlich unterschiedlich sind. Nur große Raumwellenlängen mit entsprechend geringer Auflösung werden vom Algorithmus richtig zugeordnet. Diese sind nämlich in beiden Bildern in ihrer Phasenlage identisch und liegen so, daß an Mo Positionen Intensitätsminima entstehen. Für kürzere Raumwellenlängen kann aufgrund der experimentellen Symmetrieabweichungen kein entsprechendes simuliertes Pendant gefunden werden. Daher „wählt“ der Algorithmus, wie zuvor erläutert, ein niedrig aufgelöstes Bild auf der Überfokussseite der „richtigen“ Lösung. Eine solche Situation ist an den niedrigen Kovarianzwerten um 0.9 zu erkennen. Nach der Symmetriekorrektur steigen die Kovarianzwerte merklich auf Werte um 0.99 und die ermittelten Werte für den Defokus ändern sich merklich in Richtung Unterfokus. Die Anwendung der Symmetriekorrektur verhindert also den Pseudo-Dämpfungseffekt, der bei unkorrigierten Bildern zur Nicht-Wahrnehmung verschobener aber dennoch im Bild deutlich vorhandener Details führt. Durch die Symmetriekorrektur wird die Mustererkennung also bedeutend „robuster“.

Die ermittelten Werte für die Probendicke weisen eine hohe Genauigkeit von 1.5 nm auf, was 4 Lagen der Ni_4Mo Einheitszelle entspricht. Möglicherweise ist die Genauigkeit des Verfahrens noch größer, da es relativ unwahrscheinlich ist, daß auf den vier Negativen trotz aller Bemühungen jeweils *genau* dieselben Einheitszellen gefunden werden konnten. Da durch Stufen an der Präparatober- und Unterseite durchaus lokale Dickschwankungen von der Höhe einer oder zwei Einheitszellen möglich sind, könnten bei einer denkbaren Verschiebung des Bildausschnitts um nur wenige Einheitszellen leichte Dickenunterschiede zwischen den auf den vier Bildern festgehaltenen Probenbereichen bestehen.

Es ist schwierig, eine quantitative Abschätzung über die Genauigkeit der ermittelten Defokuswerte zu machen. Sogar wenn ein amorpher Probenbereich in der Nähe der untersuchten Probenstelle gewesen wäre, hätte dieser nicht zur Bestimmung des Defokus mittels Auswertung der Thonschen Ringe [65] herangezogen werden

können. Die untersuchte Probenstelle ist mit 14 nm zu dick, um noch als schwaches Phasenobjekt zu gelten, was jedoch Voraussetzung für die Auswertung wäre. Eine Messung der relativen Defokusabstände war ebensowenig möglich, da mechanische Schwingungen, die vom Filmladesystem auf die Mikroskopsäule übertragen wurden, die vertikale Probenlage zwischen zwei Aufnahmen auf unkontrollierbare Art und Weise veränderten. Die hervorragende Übereinstimmung zwischen den bearbeiteten experimentellen Bildern und den am besten korrelierten simulierten Bildern läßt jedoch vermuten, daß die Genauigkeit der ermittelten Defokuswerte für jede denkbare weitere Anwendung ausreichend ist.

Generell hängt die Anwendbarkeit sowie die Genauigkeit der vorgestellten Methode von der Struktur des abgebildeten Materials sowie von den Transfereigenschaften des verwendeten Mikroskops ab. Im Falle von Elementhalbleitern wie kristallinem Si oder Ge als Objekt, wo aufgrund der kleinen Gitterkonstante nur eine kleine Zahl von Strahlen vom Mikroskop übertragen wird, ist die mögliche Variation des Bildkontrasts sehr beschränkt. Ähnliche oder sogar identische Bildmuster existieren hier in weit voneinander entfernten „Inseln“ der Defokus-Dicken-Ebene. Da der Bildkontrast in solch einem Fall keine eindeutige Information über Probendicke und Defokus enthält, gestattet die hier diskutierte Methode nur eine Bestimmung von möglichen Kombinationen von Defokus- und Dickenwerten, bei denen ein gegebener Bildkontrast auftreten kann. Ein anderes Problem tritt bei Strukturen mit geringer Symmetrie auf. Bei polaren Halbleitern, wie zum Beispiel GaAs oder InP, treten Kontrastmaxima oder Minima nicht unbedingt an der Stelle projizierter Atomkoordinaten oder an Tunnelpositionen zwischen Atomen auf, wenn die polare Achse in der Abbildungsebene liegt. Bei Betrachtung entlang der [110] Zonenachse können sich Kontrastmaxima oder Minima innerhalb eines bestimmten Bereiches kontinuierlich entlang der polaren Achse um die wahren Atompositionen verschieben. Da man a priori nicht wissen kann, wo die projizierten Atompositionen in einem bestimmten Bild zu liegen kommen, ist es nicht möglich, einen zwar beliebigen aber gemeinsamen Koordinatenursprung für den Vergleich mit simulierten Bildern festzulegen. Im Gegensatz zu dem in dieser Arbeit behandelten Ni_4Mo , bei dem aufgrund der hohen Symmetrie nur zwei Ursprungsvarianten entsprechend den hochsymmetrischen Punkten der Einheitszelle überprüft werden mußten, muß in solch einem Falle mit großem Aufwand ein ganzes Kontinuum von möglichen Nullpunkten beim Bildvergleich berücksichtigt werden.

Die hier vorgestellte Methode ist am besten geeignet für Materialien mit relativ großen Einheitszellen, die mindestens Zentrosymmetrie besitzen. Ein Beispiel hierfür ist der Hochtemperatursupraleiter $\text{YBa}_2\text{Cu}_3\text{O}_7$, der einerseits aufgrund der mmm-Symmetrie der Einheitszelle den Symmetrieanforderungen genügt. Andererseits ist aufgrund der mit mehreren Atomsorten besetzten Einheitszelle, die entlang einer Achse mehr als 1 nm mißt, die Forderung nach stark variablem Bildkontrast erfüllt, falls diese Achse in der Bildebene liegt.

Kapitel 8

Rekonstruktion der Austrittswellenfunktion

8.1 Einführung

8.1.1 Zielsetzung

Wie in der Einleitung der vorliegenden Arbeit anhand von Abbildung 1.1 erläutert wurde, ist der hochauflösende Abbildungsprozeß von zwei Stufen der Informationsverschlüsselung geprägt, die es erschweren, die atomare Struktur eines Objekts anhand der gewonnenen Abbildungen direkt zu interpretieren.

Der erste Verschlüsselungsprozeß ist auf die quantenmechanische Natur der Elektronenbeugung zurückzuführen. Die Behandlung des einfallenden Elektronenstrahls als Welle, die an den elektrostatischen Potentialen der Atome gebeugt wird, führt zu komplexen Interferenzerscheinungen innerhalb des Objekts. Allein dieser Vorgang kann zu einem Verlust einer 1:1 Relation zwischen der projizierten Struktur und der Amplitude und der Phase der resultierenden Wellenfunktion an der Austrittsfläche des Objekts führen.

Bei der Propagation von der Austrittsfläche des Objekts zur Bildebene wird diese Elektronenwelle abermals aufgrund der sphärischen Aberration der Objektivlinse und aufgrund der Möglichkeit zur Defokussierung auf intuitiv nicht nachvollziehbare Art und Weise in Amplitude und Phase modifiziert. Schließlich führt der Abbildungsprozeß zum Verlust der Phase der durch die Elektronenoptik modifizierten Wellenfunktion, da in die Bildintensität nur die Wellenamplitude eingeht. Jedoch auch die Amplitude der modifizierten Wellenfunktion ist nicht mehr aus der beobachteten Bildintensität analytisch ableitbar, da aufgrund der partiellen Kohärenz des einfallenden Elektronenstrahls Amplitude und Intensität nach nichtlinearen Gesetzmäßigkeiten verknüpft sind.

Da die beobachtete Bildintensität aufgrund der beschriebenen Abbildungsvorgänge nur mehr sehr indirekt mit der Struktur verknüpft ist, bietet eine Rekonstruktion der Austrittswellenfunktion eine Reihe von Vorteilen, die im folgenden erläutert werden.

Die Austrittswellenfunktion ist, abgesehen von der Wellenlänge des einfallen-

den Elektronenstrahls, frei von gerätespezifischen Parametern. Die Rekonstruktion der Wellenfunktion bietet somit die Möglichkeit, die in elektronenmikroskopischen Abbildungen enthaltenen meßtechnischen Kontrastartefakte zu eliminieren und die Objektinformation „herauszufiltern“.

Die rekonstruierte Wellenfunktion enthält Information über den gesamten Raumfrequenzbereich, der unterhalb des Informationslimits liegt. Nach dem Kriterium von Linfoot [96] ist im Falle schwacher Phasenobjekte die Qualität einer Abbildung durch die von der Kontrasttransferfunktion eingeschlossene Fläche bestimmt. Diese Fläche wird umso geringer, je mehr Nullstellen in der Kontrasttransferfunktion auftreten. Dies ist der Fall bei typischen Paßbandabbildungen. Hierbei wird die Fokussierung üblicherweise so gewählt, daß ein schmaler Raumfrequenzbereich mit hoher Auflösung auf Kosten anderer Raumfrequenzbereiche abgebildet wird. Im Gegensatz dazu erfüllt die rekonstruierte Wellenfunktion das Kriterium von Linfoot im übertragenen Sinne maximal, da sie keinerlei Transferlücken besitzt und alle Raumfrequenzen mit gleicher Gewichtung enthält.

Die rekonstruierte Phase ist bei geringen Objektdicken proportional zum Kristallpotential. Im Falle eines schwachen Phasenobjektes ist somit das letztendlich verfolgte Ziel der direkten Strukturbestimmung allein mittels der Rekonstruktion der Wellenfunktion erreicht. Eine analytische oder numerische Invertierungsprozedur des Elektronenbeugungsprozesses, die nötig wäre, um bei dickeren Präparaten bei gegebener Wellenfunktion das Kristallpotential und somit die atomare Struktur zu ermitteln, ist für den allgemeinen Fall noch nicht verfügbar. Es existieren jedoch aussichtsreiche Ansätze zur Lösung dieses Problems [97, 98, 99].

8.1.2 Bestehende Methoden

Die Rekonstruktion der Austrittswellenfunktion ist noch keine Standardprozedur in der hochauflösenden Elektronenmikroskopie und wurde erst in wenigen Fällen erfolgreich durchgeführt. Die bestehenden Rekonstruktionsmethoden können in drei Kategorien eingeteilt werden.

Auf der einen Seite wird die vielversprechende Methode der Elektronenholographie weiterentwickelt [100, 7, 8, 101, 102, 103, 104]. Diese Methode erfordert eine Modifikation eines konventionellen Elektronenmikroskops durch den Einbau eines elektrostatischen Biprismas in der Ebene des ersten Zwischenbildes. Die Methode kann nur direkt am Objektrand eingesetzt werden, wobei nur ein Teil des einfallenden Strahls das Objekt „beleuchtet“ und der andere Teil am Rand „vorbeigeht“. Das Biprisma vereinigt den Teil des einfallenden Strahls, der das Objekt ungebeugt am Rand passiert hat (Referenzstrahl), mit dem Teil des einfallenden Strahls, der in der Probe gebeugt wurde, zu einem sogenannten Interferogramm. In diesem Interferogramm sind Amplitude und Phase der Wellenfunktion der Bildebene gespeichert.

Die Austrittswellenfunktion kann durch einfache numerische Rückpropagation der Bildebenenwelle an die Objektunterseite ermittelt werden. Ein Vorteil dieser Methode besteht darin, daß die Wellenfunktion in der Bildebene direkt gemessen werden kann. Abgesehen von der Rückpropagation findet die eigentliche Rekonstruktion, nämlich die Entflechtung der bei konventionellen Aufnahmen auf komplizierte

Art „verwobenen“ Zweistrahl-Interferenzterme, schon im Experiment statt. Neben der Tatsache, daß diese Methode wegen des geringen Strahlüberlapps nur in unmittelbarer Nähe des Objektrandes angewandt werden kann, gibt es noch einen weiteren Nachteil: Aufgrund der Nutzung einer sogenannten Trägerfrequenz sind die Anforderungen bezüglich der benötigten Zahl der Bildabtastpunkte mindestens dreimal höher als bei einer konventionellen Aufnahme, was den auf einem Interferogramm festgehaltenen Bildbereich sehr einschränkt bzw. den Einsatz unpraktikabel großer Pixelfelder bei einer digitalen Bildaufnahme erforderlich machen würde [105].

Rekonstruktionsmethoden, die mit dem konventionellen experimentellen Aufbau eines Elektronenmikroskops auskommen, lassen sich je nachdem, ob in ihnen nichtlineare Interferenzerscheinungen berücksichtigt werden, in zwei Kategorien einteilen. Diese zwei Kategorien werden im folgenden beschrieben.

Lineare Rekonstruktionsmethoden [106, 107, 108] basieren meist auf der Annahme, daß die Näherung des schwachen Phasenobjektes gilt. Für sehr dünne Objekte (schwache Phasenobjekte) entspricht die Bildintensität einer Faltung des Objektpotentials mit der von der Defokussierung abhängigen Impulsantwortfunktion (engl. *point spread function*, *impulse response function*) des Mikroskops. Die Entfaltung der Bildintensität erreicht man mittels der simplen Division der Fouriertransformierten der Bildintensität durch die Kontrasttransferfunktion. Die nach Rücktransformation in den Realraum erhaltene Intensitätsverteilung ist proportional zum Objektpotential. Transferlücken, die durch Nullstellen der Kontrasttransferfunktion hervorgerufen werden, können durch die Berücksichtigung zusätzlicher, unter anderen Defokuswerten aufgenommener Bilder „geschlossen“ werden. Schwache Phasenobjekte sind in der Praxis jedoch extrem selten.

Während lineare Rekonstruktionsmethoden im Prinzip mit einem einzigen Eingabebild durchführbar sind, verknüpfen Methoden, die nichtlineare Kontrastbeiträge berücksichtigen, den Informationsgehalt mindestens zweier Bilder. Anhand analytischer Berechnungen in nur einer Dimension wurde gezeigt, daß zwei Bilder, die mit unterschiedlicher Defokussierung bei vollständig kohärenter Beleuchtung aufgenommen werden, ausreichende Information enthalten können, um die Wellenfunktion stark streuender Objekte zu rekonstruieren [109]. Misell [110] führte einen Algorithmus ein, der auf diesem Prinzip basiert. Es ist interessanterweise auch möglich, die Information aus der Bildebene mit der Information aus der Beugungsebene zu verknüpfen. Ein Algorithmus, der von Gerchberg und Saxton [111] entwickelt wurde, benötigt als Eingabedaten eine Abbildung sowie ein Beugungsbild von derselben Objektstelle. Untersuchungen zur Eindeutigkeit sowie zur physikalischen Realisierbarkeit der von diesem Algorithmus produzierten Lösungen sind äußerst komplex und konnten ebenso nur unter idealisierten Annahmen analytisch durchgeführt werden [112]. Beide Algorithmen berücksichtigen jedoch nicht den Einfluß eines partiell kohärenten Elektronenstrahls auf die Abbildung. Die partielle Kohärenz des einfallenden Elektronenstrahls ist jedoch der ursächlich auflösungsbegrenzende Faktor in der hochauflösenden Elektronenmikroskopie und kann somit für eine quantitative Rekonstruktion nicht außer Acht gelassen werden.

Kirkland [113, 114, 115] entwickelte eine rekursive Rekonstruktionsmethode, die auf der Auswertung einer ganzen Defokus-Serie beruht. Der Einfluß partiell kohären-

ter Einstrahlung wird bei dieser Methode in vollem Maße berücksichtigt. Ausgangspunkt ist eine Schätzung der gesuchten Wellenfunktion, die schon möglichst genau der „wahren“ Lösung entsprechen sollte. Durch Auswertung der Differenz zwischen den aus der geschätzten Wellenfunktion berechneten Bildern und den experimentellen Bildern wird in einer Art Rückkopplung eine neue Wellenfunktion berechnet. Diese Wellenfunktion ersetzt die erste Schätzung und der Vorgang wird so lange wiederholt, bis eine zufriedenstellende Übereinstimmung zwischen den berechneten und den experimentellen Bildern gefunden wird. Es wurde allerdings gezeigt [116], daß diese Methode nicht konvergiert, falls die Näherung des schwachen Phasenobjektes nur mehr unzureichend zutreffend ist. In diesem Fall konvergiert der Algorithmus an einem „lokalen“ Maximum der Übereinstimmung zwischen berechneten und experimentellen Bildern. Die „richtige“ Lösung, die dem globalen Maximum entspricht, wird übersehen. Ein sehr schneller rekursiver Algorithmus, der dem obigen ähnelt, wurde für spezielle Abbildungsbedingungen entwickelt, die beim Einsatz hochkohärenter Feldemissionsquellen gegeben sind. Wegen des dadurch möglichen faktorisierten Ansatzes zur Bildberechnung sind enorme Rechenzeiterparnisse möglich [10].

Eine ganz andere Möglichkeit, nichtlineare Effekte zu berücksichtigen, wird im Rahmen der von Van Dyck entwickelten „Fokusvariationsmethode“ eingeschlagen [9, 11]. Die Methode basiert auf der näherungsweisen Extraktion linearer Kontrastbeiträge aus den Bildern einer Defokus-Serie. Mit den so gewonnenen linearen Bildbeiträgen wird eine Rekonstruktion erheblich vereinfacht. Bei dieser Methode muß eine Defokus-Serie aus typischerweise 40 – 60 Bildern mit sehr kleinem Defokusinkrement ΔZ von $\Delta Z \approx 1$ nm angefertigt werden. Die Bildintensität der gesamten Serie $I(\vec{r}, Z)$ wird von dem durch $\vec{r}$ und Z aufgespannten Realraum in den dreidimensionalen Fourierraum transformiert. Das Resultat ist dann die Bildintensität $I(\vec{g}, \xi)$, wobei $\vec{g}$ ein reziproker Gittervektor und ξ die zum Defokus Z konjugierte reziproke Größe ist. Aufgrund der Periodizität der Aberrationsfunktion bezüglich Defokusänderungen [93] liegen die Fourierkoeffizienten linearer Kontrastbeiträge auf einer Parabel $\xi = \alpha g^2$. Da nichtlineare Kontrastbeiträge weiter im Fourierraum verstreut liegen, können lineare Anteile separiert werden. Bei N Eingabebildern sollte sich das Verhältnis von linearem zu nichtlinearem Bildkontrast innerhalb der gesamten Serie nach Rücktransformation in den Realraum um $\sqrt{N}$ verbessert haben. Bei dieser Methode wird der nichtlineare Bildkontrast innerhalb der Serie als Rauschen angesehen und es wird davon ausgegangen, daß der lineare Bildkontrast von vornherein dominant ist. Im Extremfall dominanten, nichtlinearen Bildkontrasts, der bei Probendicken auftreten kann, an denen die „Pendellösung“ für den Nullstrahl ein Minimum hat, ist die Methode nicht anwendbar.

In der vorliegenden Arbeit wird eine alternative Methode zur Rekonstruktion der Austrittswellenfunktion aus Bildern einer Defokus-Serie vorgestellt. Es wird gezeigt, daß nicht-deterministische Algorithmen geeignet sind, die zu periodischen Bildmotiven gehörende Austrittswellenfunktion zu ermitteln. Effekte partieller Kohärenz werden dabei in vollem Maße berücksichtigt.

8.2 Theorie

Im Prinzip kann die zu einer Defokus-Serie gehörende Wellenfunktion dadurch ermittelt werden, daß man die konventionelle Bildsimulation mit „geratenen“ Wellenfunktionen solange ausprobiert, bis die berechneten Bilder mit den experimentellen Bildern übereinstimmen. Da die Anzahl der zu überprüfenden Wellenfunktionen beinahe unendlich ist, ist eine solche Strategie ohne die Einbeziehung von Zusatzbedingungen sicherlich nicht praktikabel. Sogar im Falle typischer periodischer Objekte, wo das Wissen über die relevanten Raumfrequenzen die Anzahl von Testwellenfunktionen enorm einschränkt, würde eine Suchmethode, die auf dem Zufallsprinzip basiert, die Rechenkapazität heutiger Computer hoffnungslos überfordern.

Die Anzahl der zu überprüfenden Möglichkeiten sei anhand eines einfachen Beispiels erläutert. Die gesuchte Wellenfunktion bestehe aus nur fünf Fourierkoeffizienten. Da die Wellenfunktion beliebig skaliert werden kann, steht es frei, die Amplitude des Nullstrahls auf den Wert 1 zu setzen. Ebenso steht es frei, einen Phasenfaktor frei zu wählen, da eine Wellenfunktion aus prinzipiellen Gründen nur bis auf einen für alle Koeffizienten gemeinsamen Phasenfaktor bestimmt werden kann. Wie auch im folgenden soll hier die Phase des Nullstrahls zu 0 gewählt werden. Tatsächlich müssen von den fünf Fourierkoeffizienten also nur vier bestimmt werden. Dies bedeutet, daß je vier Amplituden- und vier Phasenwerte gesucht sind. Die Amplituden sollen auf einer Skala von 0 bis 1 in Schritten von einem $1/100$ bestimmt werden, die Phasen sollen auf dem Bereich von 0° bis 360° in Schritten von 3.6° ermittelt werden. Für jeden Fourierkoeffizienten gibt es somit $100 \times 100 = 10\,000$ Möglichkeiten. Sind die vier Strahlen über nichtlineare Interferenzen untereinander stark gekoppelt, so müssen $10\,000^4 = 10^{16}$ Möglichkeiten überprüft werden. Erfordert die Überprüfung einer Möglichkeit nur eine Millisekunde, so wäre bei einer Suche nach Zufallsprinzip im schlimmsten Falle erst nach etwa 300 000 Jahren die Lösung gefunden.

Eine effizientere Klasse von Suchverfahren, die den unpraktikabel hohen Aufwand einer Zufallssuche begrenzt, benötigt als Eingabe eine ungefähre Schätzung der zu erwartenden Lösung. Anstatt zufällig den Lösungsraum zu durchsuchen, werden Schritt für Schritt kleine Änderungen der Amplituden- und Phasenwerte in der Nähe der geschätzten Lösung untersucht. Die Methode der konjugierten Gradienten [115], die Methode des „stochastic hillclimbing“, sowie eine große Anzahl von Variationsmethoden sind Beispiele für solche *lokale* Suchtechniken. Diese Methoden besitzen eine äußerst hohe Sucheizienz. Der Erfolg solcher Methoden hängt jedoch extrem von der Qualität der anfänglichen Lösung ab. Wenn die anfänglich geschätzte Wellenfunktion zu sehr von der „wahren“ Lösung abweicht, versagen solche Suchtechniken komplett. Gerade bei der hier besprochenen Anwendung ist es äußerst schwierig, eine relevante Schätzung der anfänglichen Lösung zu machen.

Aus den obigen Ausführungen wird klar, daß effektive Suchstrategien zwei grundlegende Eigenschaften besitzen sollten. Zum einen müssen alle wichtigen Gegenden des Suchraums „erkundet“ werden. Zum anderen müssen in gewissem Maße die besten Lösungen, die dabei gefunden werden, optimiert werden. Die Forderung nach einer *globalen* Suchstrategie, die zur Erkundung des Suchraums nötig ist, sowie die Forderung nach einer *lokalen* Suchstrategie zur Optimierung bereits erworbe-

ner Kenntnisse, sind in sich widersprüchlich. Eine erfolgreiche Suchstrategie erfordert deshalb ein ausgeglichenes Verhältnis zwischen den beiden Gegenpolen *globale Exploration* und *lokale Optimierung*.

In den letzten Jahrzehnten wurden zwei wichtige stochastische Suchtechniken entwickelt, die die oben geschilderten Anforderungen erfüllen. Die Methode des *simulated annealing* sowie der Gebrauch *genetischer Algorithmen* ermöglichen das erfolgreiche Auffinden von Lösungen innerhalb sehr ausgedehnter Suchräume. Eine gute Beschreibung dieser Algorithmen, die im folgenden zur Rekonstruktion der Austrittswellenfunktion eingesetzt werden, befindet sich in der Artikelsammlung [117].

8.2.1 Evaluation von Testwellenfunktionen

Im folgenden wird die Übereinstimmung zwischen einer Testwellenfunktion und der unbekannten und noch zu suchenden experimentellen Wellenfunktion indirekt anhand der Übereinstimmung der zugehörigen Bilder bestimmt. Es wird eine Größe benötigt, die quantitativ beschreibt, wie gut eine Testwellenfunktion Ψ , die als Eingabe für die elektronenoptische Bildsimulation dient, den experimentellen Bildkontrast reproduziert. Die sogenannte Evaluationsfunktion $E(\Psi)$ wird zu diesem Zwecke als mittlere Winkelabweichung zwischen den von der Testwellenfunktion generierten Bildern und den experimentellen Bildern definiert. Im Falle von N experimentellen Bildern, die bei verschiedenen Defokuswerten Z_i mit $i = 1, \dots, N$ aufgenommen wurden, lautet die Evaluationsfunktion $E(\Psi)$:

$$E(\Psi) = \frac{1}{N} \sum_{i=1}^N \arccos(c(Z_i)). \quad (8.1)$$

Hierbei ist $c(Z_i)$ der Kreuzkovarianzkoeffizient zwischen einem von der Testwellenfunktion zu einem bestimmten Defokuswert Z_i generierten Bild und dem entsprechenden experimentellen Bild. Der Kreuzkovarianzkoeffizient wird in der in Abschnitt 7.2.2 in Gleichung (7.9) definierten Fourierraumdarstellung berechnet, da die Bildberechnung analog zu Abschnitt 7.2.1 im Fourierraum erfolgt. Der Grund, warum in Gleichung (8.1) nicht der Kreuzkovarianzkoeffizient selbst, sondern der Winkel $\Phi = \arccos(c)$ herangezogen wird, ist folgender: Falls zwei Bilder fast gleich sind, erreicht der Kreuzkovarianzkoeffizient seinen „Sättigungsbereich“ ($\cos \Phi \approx 1$ für $\Phi < 5^\circ$). Da bei der Rekonstruktion der Wellenfunktion geringe Unterschiede zwischen experimentellem und simuliertem Bild eine Rolle spielen können, ist der Winkel Φ , der in einem solchen Fall eine stärkere Steigung besitzt, vorzuziehen.

Falls in Gleichung (8.1) $E(\Psi) = 0^\circ$ ist, stimmt die Defokus-Serie, die mit der Testwellenfunktion Ψ generiert wurde, perfekt mit der experimentellen Defokus-Serie überein. $E(\Psi) = 90^\circ$ hingegen bedeutet, daß überhaupt keine Korrelation zwischen den mit der Testwellenfunktion generierten Bildern und den experimentellen Bildern besteht.

8.2.2 Der Simulated-Annealing-Algorithmus

Die Methode des *simulated annealing* basiert auf fundamentalen Prinzipien der statistischen Mechanik. Grundlegende Elemente dieser Methode wurden von Metropolis und Mitarbeitern bei der Untersuchung eines Vielteilchensystems im thermischen Gleichgewicht entwickelt [118]. Die Methode wurde von Kirkpatrick und Mitarbeitern im Rahmen der Behandlung von kombinatorischen Optimierungsproblemen weiterentwickelt [119, 120].

Das Prinzip des *simulated annealing* basiert auf der Beobachtung, daß ein Vielteilchensystem im thermischen Gleichgewicht bei gegebener Temperatur T ein Minimum der Freien Energie einnimmt. Die Suche nach einem globalen Minimum der Evaluationsfunktion $E(\Psi)$ aus Gleichung (8.1), die sich auf ein N -Strahlen Problem bezieht, kann analog zur Suche nach einem globalen Minimum der Freien Energie eines Systems aus N wechselwirkenden Teilchen behandelt werden.

In der statistischen Mechanik ist die Wahrscheinlichkeit π_s , im thermischen Gleichgewicht ein System in einer bestimmten Konfiguration s zu finden, durch die Boltzmann-Verteilung gegeben:

$$\pi_s = \frac{\exp\left(-\frac{E_s}{kT}\right)}{\sum_{w \in S} \exp\left(-\frac{E_w}{kT}\right)}. \quad (8.2)$$

In der obigen Gleichung bezeichnet S die Gesamtheit aller möglichen Konfigurationen, E bezeichnet die zu einer bestimmten Konfiguration gehörende Konfigurationsenergie, T bezeichnet die Temperatur und k ist die Boltzmannkonstante. Der Übergang von einer bestehenden Konfiguration i zu einer alternativen Konfiguration f tritt dabei mit einer Wahrscheinlichkeit $p_{i \rightarrow f}$ auf, wobei

$$p_{i \rightarrow f} = \frac{\pi_f}{\pi_i} = \exp\left(-\frac{E_f - E_i}{kT}\right). \quad (8.3)$$

Der Übergang erfolgt immer, falls $E_f \leq E_i$ ist, sonst erfolgt er mit einer Wahrscheinlichkeit $p < 1$. Das thermodynamische System befindet sich im Konfigurationsraum auf einer Schale, deren Breite proportional zu kT ist. Eine Senkung der Temperatur führt zu einer Verkleinerung der Oberfläche als auch der Breite dieser Schale. Im Grenzfall $T \rightarrow 0$ nimmt das System entsprechend der niedrigst möglichen Konfigurationsenergie nur mehr wenige oder gar nur mehr eine einzige Konfiguration ein. Solch eine Konfiguration wird Grundzustand genannt.

Der Simulated-Annealing-Algorithmus beinhaltet folgende Komponenten: Eine Darstellung der zu optimierenden Konfigurationen, einen Generator für die Startkonfiguration, einen Generator für neue Konfigurationen, eine Evaluationsfunktion E , die eine aktuelle Konfiguration bewertet, eine Funktion, die die Übergangswahrscheinlichkeit $p_{i \rightarrow f}$ berechnet, sowie einen „Temperaturfahrplan“, der den Verlauf des Abfalls von der Ausgangstemperatur zu einer Endtemperatur nahe Null beschreibt. Da die „Energie“ beziehungsweise die Evaluationsfunktion aus Gleichung (8.1) dimensionslos ist, wird der Faktor kT aus Gleichung (8.3) entsprechend durch eine dimensionslose Größe ersetzt.

```
Setze Anfangstemperatur;  
Generiere Anfangskonfiguration;  
Berechne Wert der Evaluationsfunktion für Anfangskonfiguration;  
Wiederhole  
  Wiederhole  
    Generiere neue Konfiguration;  
    Berechne Wert der Evaluationsfunktion für neue Konfiguration;  
    Berechne Übergangswahrscheinlichkeit zur neuen Konfiguration;  
    Entscheide, ob neue Konfiguration akzeptiert wird;  
    Falls akzeptiert:  
      Ersetze aktuelle Konfiguration durch neue  
bis thermisches Gleichgewicht erreicht ist;  
Erniedrige Temperatur um kleinen Betrag  
bis Grundzustand erreicht ist.
```

Abbildung 8.1: Kompakte Darstellung des Simulated-Annealing-Algorithmus

Im Rahmen der hier vorgestellten Anwendung des Algorithmus ist die zu optimierende Konfiguration ein Satz von Amplituden- und Phasenwerten, die die Fourierkoeffizienten $F(\vec{k})$ der Testwellenfunktion repräsentieren (Gl. (7.1) in Abschnitt 7.2.1.). Eine Startkonfiguration aus Amplituden- und Phasenwerten wird mit Hilfe eines Zufallszahlengenerators erstellt. Eine neue Konfiguration wird dadurch erstellt, daß entweder der Amplituden- oder der Phasenwert eines mit Hilfe des Zufallszahlengenerators ausgewählten Strahls durch eine Zufallszahl ersetzt wird. Die „Energie“ der anfänglichen sowie der so geänderten Wellenfunktion wird mittels der in Gleichung (8.1) dargestellten Evaluationsfunktion ermittelt. Die Übergangswahrscheinlichkeit von der anfänglichen Konfiguration zur geänderten Konfiguration wird mittels Gleichung (8.3) berechnet. Der Übergang wird akzeptiert, falls die Wahrscheinlichkeit größer als eine zwischen 0 und 1 ermittelte Zufallszahl ist, sonst wird der Übergang abgelehnt. Nach einer bestimmten Anzahl akzeptierter Übergänge wird von der aktuellen Konfiguration angenommen, daß sie einen Gleichgewichtszustand repräsentiert, und die Temperatur wird nach einer Vorschrift $T_{n+1} = cT_n$ mit $c < 1$ gesenkt. Dieser Vorgang wird solange wiederholt bis die Temperatur einen Wert nahe Null erreicht. Der logische Ablauf des Simulated-Annealing-Algorithmus ist in kompakter Form in Abbildung 8.1 dargestellt.

Der Simulated-Annealing-Algorithmus ist nicht sehr effizient, wenn es darum geht, Amplituden und Phasen, die bereits sehr nahe an der gesuchten Lösung sind, weiter zu optimieren. Obwohl die Lösung „fast“ gefunden ist, wird die globale Suche unnötigerweise bis zum Ende des Optimierungsprozesses fortgesetzt. Dies führt zu einer übermäßigen Verschwendung von Rechenaufwand. Eine Möglichkeit, diesen Aufwand zu vermeiden, ist der Einsatz lokaler deterministischer Suchalgorithmen, die den Optimierungsprozeß im späten Stadium der Suche effizienter bewältigen. Es ist jedoch sehr schwierig, den Zeitpunkt des Einsatzes eines lokalen Suchalgorithmus von vornherein festzulegen, da nicht bekannt ist, ob noch mit lokalen Mi-

nima zu rechnen ist, oder ob diese bereits vollständig überwunden sind. Eine andere Lösungsmöglichkeit, die in der gegenwärtigen Implementation der Programme genutzt wird, ist der Einbau lokaler Suchcharakteristiken in den bestehenden Algorithmus. Dazu werden parallel zu den ohnehin stattfindenden Konfigurationswechseln noch zusätzliche Konfigurationswechsel durchgeführt, die nur minimalen Änderungen der aktuellen Amplituden- und Phasenwerte entsprechen. Durch diese zusätzlichen Konfigurationswechsel wird die Konvergenz des Algorithmus gegen Ende des Optimierungsprozesses enorm beschleunigt. Da diese zusätzlichen Konfigurationswechsel parallel zu den regulären durchgeführt werden, wird die globale Suche nicht beeinträchtigt und der Algorithmus in seiner Allgemeinheit nicht eingeschränkt.

8.2.3 Genetische Algorithmen

Die Funktionsweise genetischer Algorithmen beruht auf einem vereinfachten Modell der Evolution, die in gewissem Sinne ein in der belebten Natur stattfindender Suchprozeß ist. Das Ziel dieses Suchprozesses ist die Optimierung einer Art in Bezug auf ihre Überlebensfähigkeit in einer vorgegebenen Umgebung. Genetische Algorithmen wurden erstmals von Holland [121] entwickelt, der zeigte, daß mit Hilfe solcher Algorithmen nahezu optimale Problemlösungen gefunden werden können.

Ein einfaches Modell der Evolution, das zur Lösung von Optimierungsproblemen ausreicht, umfaßt die folgenden Komponenten: Eine chromosomenähnliche Repräsentation der zu optimierenden Art, einen Satz von Individuen innerhalb dieser Art, wobei jedes Individuum eine unterschiedliche Chromosomenkonfiguration besitzt, eine Evaluationsfunktion, die die Überlebensfähigkeit eines Individuums in Bezug auf die Umgebung feststellt, sowie die genetischen Operatoren *Mutation* und *Paarung* [117].

In genetischen Algorithmen bewirkt der Mutationsoperator eine zufällige Änderung eines einzelnen Gens oder einer Gruppe von Genen innerhalb eines Chromosoms (siehe Abb. 8.2). Die Paarung wird oft dadurch realisiert, daß eine Gensequenz zwischen zwei Chromosomen ausgetauscht wird. Dabei werden Eigenschaften zweier Individuen auf neue Art kombiniert. Die Wahrscheinlichkeit, daß ein Individuum für die Paarung ausgesucht wird, ist dabei umso höher, je besser seine Anpassung an die Umwelt ist.

Die Effizienz genetischer Algorithmen beruht auf der Paarungsoperation, da es wahrscheinlich ist, daß die Paarung zweier „guter“ Individuen zu Nachkommen noch höherer Effizienz führt. Die Suche nach der bestmöglichen Anpassung an die Umwelt findet hier auf implizit parallele Art und Weise statt, da die Vielfalt aller Individuen genutzt wird, um die gesamte Population zu optimieren. Besonders geeignete Gensequenzen werden dabei zu Bausteinen für neue Chromosomen und werden als Wissensbasis von Generation von Generation weitervererbt. Ohne Anwendung des Mutationsoperators würde es jedoch zu einer Angleichung aller Individuen kommen, da durch die vielfache Paarung die in der Anfangspopulation vorhandenen Gene sehr schnell optimal kombiniert und weiterkopiert werden. Dies führt zu Inzuchterscheinungen und der Evolutionsprozeß käme zum Erliegen. Der Mutationsoperator verhindert Inzuchterscheinungen indem er den Genpool laufend „auffrischt“.

Genetischer Algorithmus

Mutationsoperator:

$$\begin{array}{c}
 (A_1, \Phi_1, A_2, \Phi_2, A_3, \Phi_3, A_4, \Phi_4) \\
 \downarrow \\
 (A_1, \Phi_1, A_2, \Phi_2, a_3, \Phi_3, A_4, \Phi_4)
 \end{array}$$

Paarungsoperator:

$$\left. \begin{array}{c}
 (A_1, \Phi_1, \boxed{A_2, \Phi_2, A_3, \Phi_3}, A_4, \Phi_4) \\
 \updownarrow \\
 (a_1, \phi_1, \boxed{a_2, \phi_2, a_3, \phi_3}, a_4, \phi_4)
 \end{array} \right\} \text{Eltern}$$

$$(A_1, \Phi_1, a_2, \phi_2, a_3, \phi_3, A_4, \Phi_4) \quad \text{Nachkomme}$$

Abbildung 8.2: Mutationsoperator und Paarungsoperator des genetischen Algorithmus. Das Chromosom eines Individuums besteht aus einer Sequenz von Genen, hier Amplitudenwerte A_i und Phasenwerte Φ_i . Oben: Bei der Mutation wird der Wert eines zufällig herausgegriffenen Gens durch einen neuen, ebenfalls durch Zufall bestimmten Wert ersetzt. Im Beispiel wird der Amplitudenwert A_3 durch den Amplitudenwert a_3 ersetzt. Unten: Bei der Paarung werden Gensequenzen (hier fett umrandet) zwischen den Chromosomen der Eltern ausgetauscht. Die so neu entstandene Gensequenz ergibt das Chromosom des Nachkommen.

```

Generiere anfängliche Population von Individuen;
Bewerte die Mitglieder der Population mittels der
Evaluationsfunktion;
Wiederhole
    Mit einer Wahrscheinlichkeit  $W_p$ :
        Wähle zwei Mitglieder der Population aus;
        Führe d. Paarungsoperation mit diesen zwei Mitgliedern durch;
        Bewerte den Nachkommen;
        Ersetze ein Mitglied der Population durch den Nachkommen
    Mit einer Wahrscheinlichkeit  $W_m$ :
        Wähle ein Mitglied aus der Population aus;
        Wähle ein Gen im Chromosom des Kandidaten;
        Ändere das Gen zufällig;
        Bewerte den Kandidaten neu
bis die Population optimiert ist.

```

Abbildung 8.3: Kompakte Darstellung eines genetischen Algorithmus

In dem zum Zwecke der Wellenrekonstruktion entwickelten genetischen Algorithmus bestehen die Chromosomen aus einer Folge von Amplituden- und Phasenwerten, die die Fourierkoeffizienten $F(\vec{k})$ der Wellenfunktion $\Psi(\vec{r})$ darstellen (Siehe Gl. (7.1) in Abschnitt 7.2.1). Eine anfängliche Population wird dadurch generiert, daß die Chromosomen aller Individuen mit gleichverteilten Zufallszahlen belegt werden. Die Mutation wird in ähnlicher Weise ausgeführt wie der Konfigurationswechsel beim Simulated-Annealing-Algorithmus. Gleichverteilte Zufallszahlen werden benutzt, um ein bestimmtes Individuum auszuwählen, um ein Gen innerhalb des Chromosoms zu lokalisieren und um den darin gespeicherten Amplituden- oder Phasenwert zu ersetzen. Die Paarungsoperation ist ebenfalls ausschließlich von Zufallszahlen bestimmt. Der erste Elternteil ist das jeweils „beste“ Individuum innerhalb der Population, der zweite Elternteil wird aus der Population durch Zufall bestimmt. Die Sequenz von Amplituden- und Phasenwerten, die zwischen den zwei Elternteilen ausgetauscht wird, kann von nur einem Gen bis zur Länge des gesamten Chromosoms variieren. Sowohl die Länge als auch die Endpunkte dieser Sequenz werden durch Zufallszahlen bestimmt. Um die Anzahl der Individuen innerhalb der Population konstant zu halten, wird das „schlechteste“ Individuum innerhalb der Population durch den bei der Paarung erzeugten Nachkommen ersetzt. Da es viele Möglichkeiten gibt, genetische Algorithmen zu verwirklichen [117], ist die kompakte Beschreibung in Abbildung 8.3 in etwas allgemeinerer Form gehalten.

Ähnlich dem Simulated-Annealing-Algorithmus ist der genetische Algorithmus ineffizient beim Optimieren bereits nahezu perfekter Lösungen. Um die Konvergenzgeschwindigkeit zu erhöhen, wurde eine ähnliche Strategie wie für das *simulated annealing* gewählt. Die Konvergenz wird beschleunigt, indem zusätzlich zu den ohnehin stattfindenden Mutationen weitere Mutationen ausgeführt werden, die um die bereits nahezu perfekten Amplituden- und Phasenwerte zentriert sind.

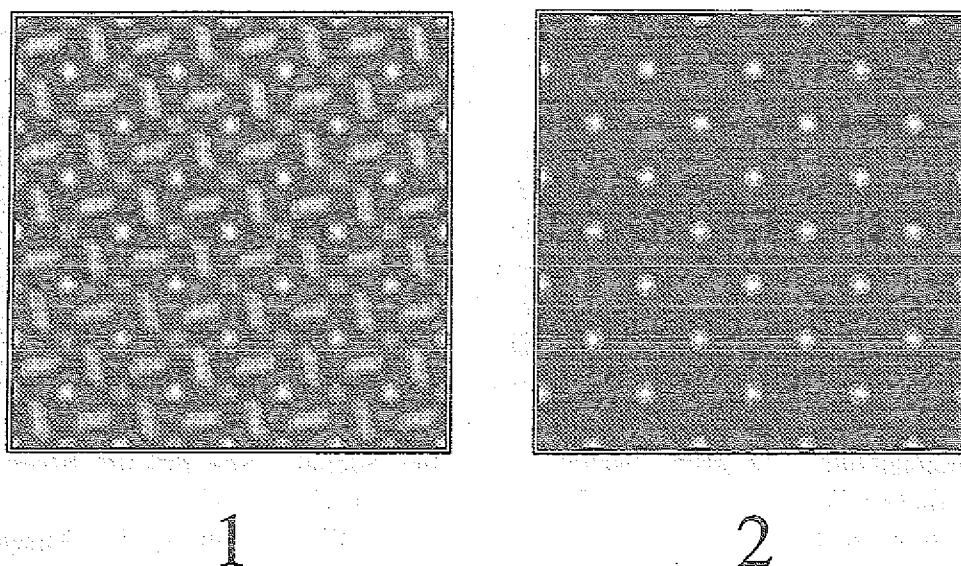


Abbildung 8.4: Simulierte Hochauflösungsbilder der Legierung Ni_4Mo . Blickrichtung ist die $[001]$ Zonenachse, jeweils 4×4 Einheitszellen sind dargestellt. Bild 1 wurde für eine Objektdicke von 25 nm und eine Defokussierung von -35 nm simuliert, die Simulation von Bild 2 basiert auf derselben Objektdicke und einer Defokussierung von -70 nm. Die Ecken der Bildberandungen entsprechen Molybdänpositionen.

8.3 Rekonstruktion der Austrittswellenfunktion aus simulierten Abbildungen von Ni_4Mo

Um die Konvergenzeigenschaften und die Genauigkeit der im Theorieabschnitt erläuterten Rekonstruktionsalgorithmen zu untersuchen, werden zwei für die $[001]$ Zonenachse simulierte Abbildungen von Ni_4Mo als Testobjekte herangezogen. Die aus jeweils 4×4 Einheitszellen bestehenden Bilder sind in Abbildung 8.4 zu sehen. Für die Simulation von Bild 1 wurde ein Defokuswert von -35 nm gewählt, Bild 2 entspricht einer Defokussierung von -70 nm. Die zu beiden Bildern gehörende Objektdicke beträgt 25 nm. Die Simulationen basieren auf einer Beschleunigungsspannung von 400 kV sowie auf einer sphärischen Aberrationskonstante von $C_s = 1$ mm. Für die Defokusstreuung wurde ein Wert von 10 nm und für den Semikonvergenzwinkel ein Wert von 0.9 mrad benutzt. Der zur Simulation verwendete Blendendurchmesser entspricht 15 nm^{-1} . Bei diesem Durchmesser liegen 29 Strahlen der $[001]$ Zonenachse von Ni_4Mo innerhalb der Objektiveblende (Abb. 3.7). Die relativ hohe Objektdicke wurde gewählt, um den Gültigkeitsbereich der Näherung des *schwachen Phasenobjekts* eindeutig zu verlassen.

8.3.1 Rekonstruktion mittels des Simulated-Annealing-Algorithmus

In diesem Abschnitt wird der Simulated-Annealing-Algorithmus auf zwei unterschiedliche Arten zur Rekonstruktion der Austrittswellenfunktion angewandt. Im ersten Fall wird das Wissen um die vierzählige Symmetrie der Ni_4Mo Struktur in [001] Projektion ausgenutzt. Jeweils vier symmetrieverwandte Strahlen werden dabei als Gruppe mit identischen Amplituden- und Phasenwerten behandelt. Anstatt der 29 Strahlen, die innerhalb der Objektivblende liegen, müssen dann nur mehr 8 ($=28/4+1$) Strahlen vom Algorithmus unabhängig variiert werden. Bei der zweiten Anwendung des Algorithmus wird keinerlei Vorteil aus dem Wissen um die Symmetrie gezogen und alle 29 Strahlen innerhalb der Objektivapertur werden unabhängig behandelt. Diese Variante ist rechenzeitaufwendiger, da der entsprechende Suchraum größer als im ersten Fall ist. Durch den Vergleich der für beide Varianten benötigten Rechenzeit läßt sich grob abschätzen wie der Suchaufwand mit der Strahlzahl ansteigt, da der Bildberechnungsaufwand in beiden Fällen gleich ist.

In Abbildung 8.5a sind Amplitude (oben) und Phase (unten) der Ausgangswellenfunktion, die benutzt wurde, um die in Abbildung 8.4 gezeigten Bilder zu generieren, dargestellt. Im folgenden wird diese Wellenfunktion als die *erzeugende* Wellenfunktion bezeichnet. Jedes der in Abbildung 8.5 gezeigten Bilder umfaßt 4×4 Einheitszellen. Während das in Abbildung 8.5 gezeigte Amplitudenmuster nicht im Zusammenhang mit der atomaren Konfiguration steht, ist das Phasenmuster eng mit der entlang [001] projizierten Struktur von Ni_4Mo verbunden. Die dunklen und scharf begrenzten Bereiche, die niedrigen Phasenwerten entsprechen, liegen an den projizierten Positionen der Nickelatome. Die typische Abfolge von jeweils 4 Nickelatomen und einem Molybdänatom entlang den $\langle 100 \rangle$ Kristallrichtungen der kfz Matrix, die in den Ausgangsbildern aus Abbildung 8.4 nicht zu erkennen ist, ist in der Phase der Austrittswellenfunktion in Abbildung 8.5 klar zu sehen.

Die Wellenfunktion, die aus den in Abbildung 8.4 gezeigten Eingabebildern mittels des Simulated-Annealing-Algorithmus im 29-Strahl Modus rekonstruiert wurde, ist in Abbildung 8.5b dargestellt. Die im 8-Strahl Modus rekonstruierte Wellenfunktion ist optisch identisch und wird deshalb nicht zusätzlich gezeigt. Durch visuellen Vergleich der Abbildungen 8.5a und 8.5b ist kein Unterschied zwischen der generierenden Wellenfunktion und der rekonstruierten Wellenfunktion zu erkennen. Die kleinen und nur mehr numerisch feststellbaren Unterschiede zwischen Eingabe- und Ausgabewellenfunktion sind in Tabelle 8.1 für die 8-Strahl Lösung aufgelistet. Abgesehen von den $\{400\}$ und den $\{330\}$ Strahlen sind die relativen Fehler der rekonstruierten Amplituden typischerweise kleiner als 0.1 % und die absoluten Fehler der Phasen sind meist kleiner als 0.1° .

Das Konvergenzverhalten des Simulated-Annealing-Algorithmus ist in Abbildung 8.6 visuell dargestellt. Die doppelt logarithmischen Diagramme zeigen den Zusammenhang zwischen der Evaluationsfunktion $E(\Psi)$ und dem Temperaturparameter T während eines typischen Programmlaufs. Das obere Diagramm in Abbildung 8.6 bezieht sich auf einen Programmlauf im 8-Strahl Modus, das untere auf einen Programmlauf im 29-Strahl Modus. Aus den Diagrammen kann entnommen

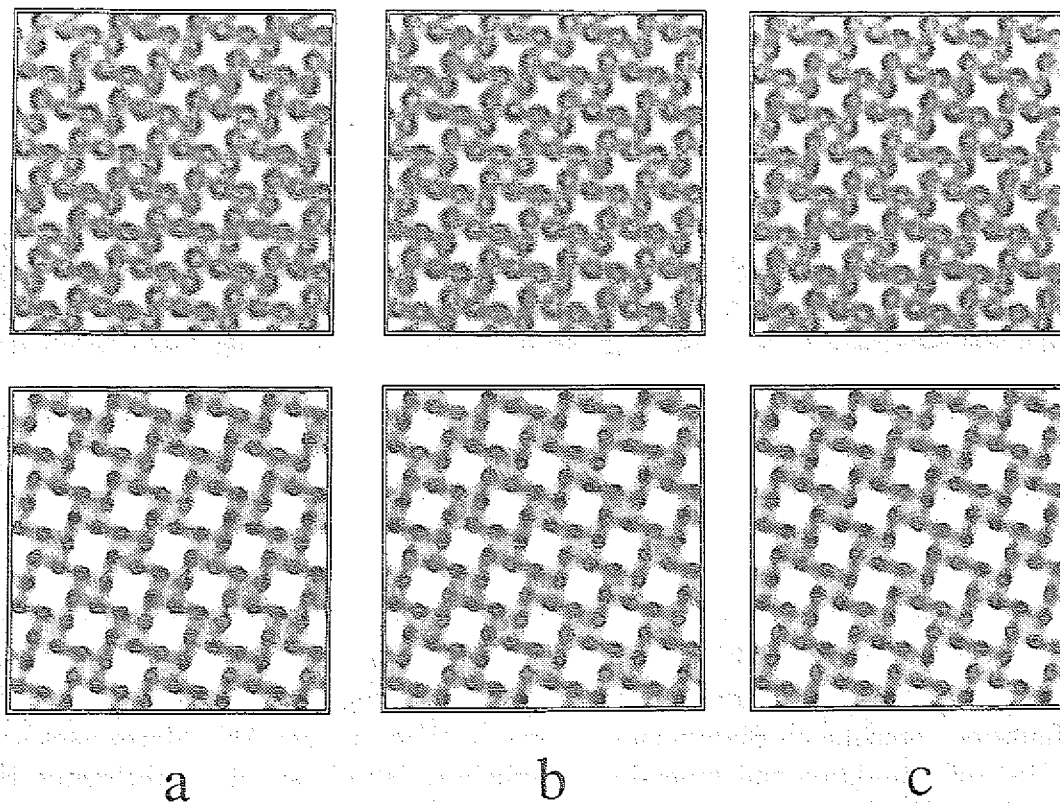


Abbildung 8.5: Austrittswellenfunktion der Legierung Ni_4Mo bei einer Objektdicke von 25 nm. Blickrichtung ist die $[001]$ Zonenachse, jeweils 4×4 Einheitszellen sind dargestellt. Die Bilder der oberen Reihe zeigen die Amplitude, die der unteren Reihe die Phase der Wellenfunktion. Für die graphische Darstellung wurde der gesamte zur Verfügung stehende Bereich einer 256-stufigen Grauskala benutzt. Die Phasendarstellung ist so normiert, daß im Fourierraum die Phase des Nullstrahls den Wert Null annimmt. (a) Ausgangswellenfunktion, die benutzt wurde, um die in Abbildung 8.4 gezeigten Testbilder zu generieren. (b) Wellenfunktion, die mittels des Simulated-Annealing-Algorithmus im 29-Strahl Modus aus den Testbildern rekonstruiert wurde. (c) Wellenfunktion, die mittels des genetischen Algorithmus im 29-Strahl Modus aus den Testbildern rekonstruiert wurde.

(hkl)	A_i	A_o	$\Delta A/A_i$ [%]	φ_i [Grad]	φ_o [Grad]	$\varphi_o - \varphi_i$ [Grad]
(000)	1.0000	1.0000	0.00	0.00	0.00	0.00
(110)	0.3341	0.3341	0.00	83.54	83.55	0.01
(200)	0.2481	0.2482	0.04	63.43	63.43	0.00
(220)	0.1287	0.1286	-0.08	30.48	30.48	0.00
(310)	0.1062	0.1062	0.00	25.38	25.38	0.00
(130)	0.2542	0.2542	0.00	223.99	224.00	0.01
(400)	0.0386	0.0384	-0.56	227.24	226.93	-0.31
(330)	0.0111	0.0103	-7.10	233.71	223.79	-9.92

Tabelle 8.1: Vergleich zwischen der erzeugenden Wellenfunktion, die benutzt wurde, um die Eingabebilder aus Abbildung 8.4 zu simulieren, und der aus diesen Bildern mittels des Simulated-Annealing-Algorithmus rekonstruierten Wellenfunktion. Die Amplituden und Phasen der „Eingabe“-Wellenfunktion sind mit dem Subskript „i“ gekennzeichnet, diejenigen der rekonstruierten Wellenfunktion mit dem Subskript „o“. Die Amplitude des Nullstrahls ist auf den Wert 1 normiert, der allen Strahlen gemeinsame Phasenfaktor wurde so gewählt, daß die Phase des Nullstrahls den Wert 0 hat.

werden, daß zu Anfang des Abkühlprozesses die Evaluationsfunktion Werte nahe 90° annimmt. In diesen hohen Werten kommt das vollkommene Fehlen einer Korrelation zwischen den Eingabebildern und den aus einer zufällig „ausgewürfelten“ Testwellenfunktion berechneten Bildern zum Ausdruck. Während des Abkühlprozesses wird die Testwellenfunktion zunehmend so modifiziert, daß die aus ihr berechneten Bilder mehr und mehr den Eingabebildern gleichen. Der Abfall der Evaluationsfunktion endet nahe 10^{-2} Grad, da die in den Programmen benutzte achtstellige Zahlengenauigkeit für die Kreuzkovarianz (Gl. (7.9)) nicht mehr ausreicht, den Unterschied zwischen den Eingabebildern aus Abbildung 8.4 und den aus der „Testwellenfunktion“ berechneten Bildern noch aufzulösen. Die aus dem Abkühlprozeß resultierende Wellenfunktion reproduziert die Eingabebilder also perfekt.

In Abbildung 8.6 ist zu sehen, daß die Fluktuationen der Evaluationsfunktion im Falle des 8-Strahl Modus erheblich größer als im 29-Strahl Modus sind. Dies liegt daran, daß die gemeinsame und gleichsinnige Modifikation von vier Strahlen, wie sie im 8-Strahl Modus erfolgt, eine größere Änderung der Evaluationsfunktion hervorruft als die Modifikation eines einzelnen Strahls im 29-Strahl Modus. Aus beiden Konvergenzdiagrammen in Abbildung 8.6 erkennt man den Zusammenhang $E(\Psi) \sim T^\alpha$ mit $\alpha \approx 1$. Obwohl die Evaluationsfunktion in den beiden Diagrammen der Abbildung 8.6 mit sinkender Temperatur etwa gleich stark abfällt, unterscheidet sich die zum Erreichen der numerisch vollständigen Konvergenz aufgewandte Rechenzeit in den beiden Fällen beträchtlich. Auf einer DEC 3500 ALPHA Rechanlage beträgt die zum Erreichen der Zahlengenauigkeitsgrenze benötigte CPU-Zeit im 8-Strahl Modus typischerweise 1–2 Minuten. Im 29-Strahl Modus sind für die Rekonstruktion derselben Wellenfunktion etwa 10 Minuten CPU-Zeit erforderlich.

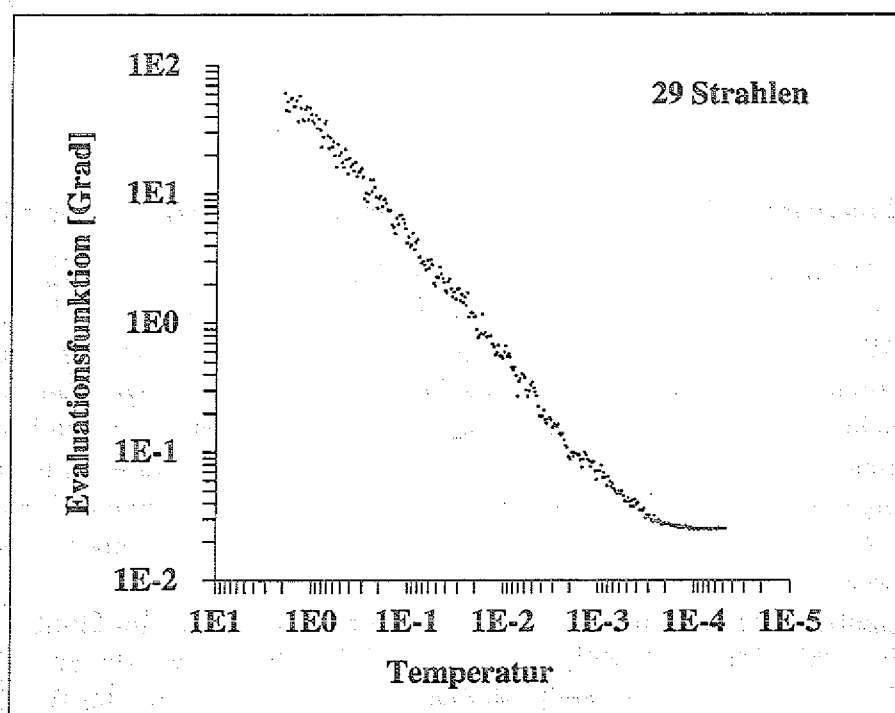
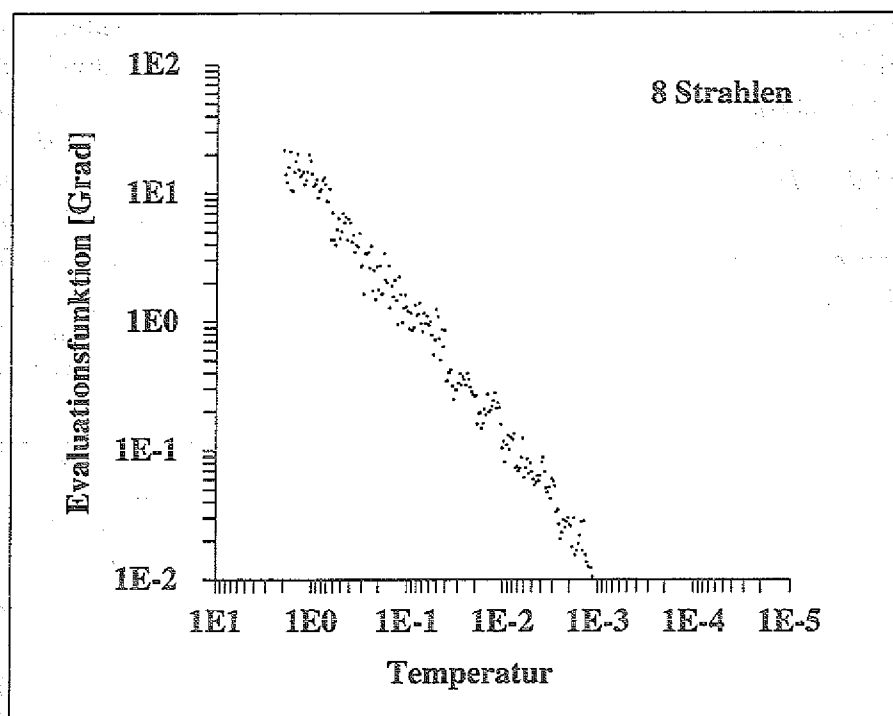


Abbildung 8.6: Konvergenzdarstellungen, die typische Programmläufe des Simulated-Annealing-Algorithmus zeigen. Oben: 8-Strahl Modus, unten: 29-Strahl Modus. Der Logarithmus der Konfigurationsenergie ist über dem Logarithmus des Temperaturparameters aufgetragen.

Um eine vorzeitige Konvergenz in einem lokalen Minimum der Evaluationsfunktion zu verhindern, kann der Abfall der Temperatur T nicht beliebig schnell erfolgen. Analog zur Metallurgie entsprechen solche lokale Minima metastabilen Zuständen, die sich nicht im thermodynamischen Gleichgewicht befinden. Die Realisierung eines metastabilen Zustands ist jedoch für die gegenwärtige Anwendung höchst unerwünscht, da ein Übergang zum Grundzustand, dem die gesuchte Wellenfunktion entspricht, ab einer gewissen Temperatur nicht mehr möglich ist. Um ein „Einfrieren“ des Algorithmus in einem lokalen Minimum zu vermeiden, muß im 29-Strahl Modus eine geringere Abkühlrate gewählt werden als im 8-Strahl Modus. Die unterschiedlichen „Temperaturfahrpläne“ spiegeln sich in den benötigten Rechenzeiten wider.

Zusätzlich wurde untersucht, ob eine Rekonstruktion der Wellenfunktion auch möglich ist, wenn nur *eines* der Bilder aus Abbildung 8.4 als Eingabe für den Algorithmus benutzt wird. Da in Bild 2 die höheren Raumfrequenzen stärker betont sind, wurde dieses Bild ausgewählt. Erstaunlicherweise war es mit dem Simulated-Annealing-Algorithmus im 8-Strahl Modus möglich, die Wellenfunktion mit der selben Genauigkeit zu rekonstruieren wie im Falle der auf zwei Bildern basierenden Rekonstruktion. Die für die Rekonstruktion aus einem Bild benötigte Rechenzeit ist jedoch etwa zehnmal länger als die für die Rekonstruktion aus beiden Bildern benötigte Rechenzeit. Aufgrund eines häufig beobachteten „Einfrierens“ des Algorithmus, muß eine sehr moderate Abkühlrate gewählt werden. Die Frage, wieviele Bilder prinzipiell zur Rekonstruktion nötig sind, wird in Abschnitt 8.5 näher behandelt.

8.3.2 Rekonstruktion mittels des genetischen Algorithmus

Um die zur Rekonstruktion der Austrittswellenfunktion eingesetzten Algorithmen vergleichen zu können, wurden die in Abbildung 8.4 gezeigten Testbilder, deren zugehörige Wellenfunktion mittels des Simulated-Annealing-Algorithmus rekonstruiert wurde, auch als Eingabe für den genetischen Algorithmus herangezogen. Wie bei der Untersuchung des Simulated-Annealing-Algorithmus wurden zwei Rekonstruktionsstrategien angewandt. Eine weniger rechenzeitaufwendige Strategie, bei der die Symmetrieeigenschaften des Materials vorab gegeben sind, sowie eine aufwendigere, bei der alle 29 Strahlen innerhalb der Objektivblende als unabhängig voneinander betrachtet werden.

Die Wellenfunktion, die mittels des genetischen Algorithmus im 29-Strahl Modus aus den in Abbildung 8.4 gezeigten Bildern rekonstruiert wurde, ist in Abbildung 8.5c dargestellt. Die im 8-Strahl Modus erhaltene Lösung für die Wellenfunktion ist nicht gesondert dargestellt, da sie optisch identisch mit der im 29-Strahl Modus ermittelten Lösung ist. Der Unterschied zwischen der erzeugenden Wellenfunktion (Abb. 8.5a) und der rekonstruierten Wellenfunktion (Abb. 8.5c) ist optisch nicht wahrnehmbar. In Tabelle 8.2 werden die Amplituden- und Phasenwerte der generierenden Wellenfunktion den Amplituden- und Phasenwerten der mittels des genetischen Algorithmus rekonstruierten Wellenfunktion gegenübergestellt. Abgesehen von den {400} und {330} Strahlen sind die relativen Fehler der Amplitudenwerte

(hkl)	A_i	A_o	$\Delta A/A_i$ [%]	φ_i [Grad]	φ_o [Grad]	$\varphi_o - \varphi_i$ [Grad]
(000)	1.0000	1.0000	0.00	0.00	0.00	0.00
(110)	0.3341	0.3340	-0.03	83.54	83.52	-0.02
(200)	0.2481	0.2479	-0.06	63.43	63.43	0.00
(220)	0.1287	0.1289	0.15	30.48	30.50	0.02
(310)	0.1062	0.1063	0.04	25.38	25.39	0.01
(130)	0.2542	0.2540	-0.06	223.99	223.98	-0.01
(400)	0.0386	0.0399	3.36	227.24	227.59	0.35
(330)	0.0111	0.0070	-36.77	233.71	240.32	6.61

Tabelle 8.2: Vergleich zwischen der erzeugenden Wellenfunktion, die benutzt wurde, um die Eingabebilder aus Abbildung 8.4 zu simulieren, und der aus diesen Bildern mittels des genetischen Algorithmus rekonstruierten Wellenfunktion. Die Amplituden und Phasen der „Eingabe“-Wellenfunktion sind mit dem Subskript „i“ gekennzeichnet, diejenigen der rekonstruierten Wellenfunktion mit dem Subskript „o“. Die Amplitude des Nullstrahls ist auf den Wert 1 normiert, der allen Strahlen gemeinsame Phasenfaktor wurde so gewählt, daß die Phase des Nullstrahls den Wert 0 hat.

in den meisten Fällen kleiner als 0.1 %. Die absoluten Fehler der Phasenwerte sind typischerweise kleiner als 0.1°. Die Lösung des 8-Strahl Problems benötigt auf einer DEC 3500 ALPHA Rechenanlage 1 – 2 Minuten an CPU-Zeit, die des 29-Strahl Problems etwa 10 Minuten.

In den in Abbildung 8.7 gezeigten Konvergenzdarstellungen, die typische Programmläufe wiedergeben, ist der zum jeweils „besten“ Individuum der Population gehörende Wert der Evaluationsfunktion logarithmisch über der Anzahl der Generationen aufgetragen. Die Punkte in den Konvergenzdarstellungen haben einen zeitlichen Abstand von jeweils 100 Generationen. Ein charakteristisches Merkmal des hier eingesetzten genetischen Algorithmus ist der anfänglich äußerst schnelle Abfall der Evaluationsfunktion. Im 8-Strahl Modus ist der Abfall während der ersten 100 Generationen so schnell, daß in der Konvergenzdarstellung nur ein Punkt über 20° zu liegen kommt. Die Konvergenzdarstellung für den 29-Strahl Modus weist ebenfalls eine äußerst starke negative Anfangssteigung auf. Solch ein „gieriges“ Verhalten des Algorithmus, bei dem die Evaluationsfunktion innerhalb weniger Generationen sogar unter 1° fallen kann, wird durch den Gebrauch niedriger Mutationsraten gefördert. In diesem Fall ist der Aufwand für die lokale Optimierung viel größer als der für die globale Erkundung des Lösungsraumes und es besteht die Gefahr, daß der Algorithmus in einem lokalen Minimum „gefangen“ wird. Ein „gutes“ Individuum der Population dominiert dann den Vermehrungsprozeß und erzeugt Nachkommen, die ihm gleichen. Schließlich gleichen sich alle Mitglieder der Population und stellen eine lokale Minimumskonfiguration dar. Bei angepaßter Mutationsrate bilden die Mitglieder der Population hingegen ein kollektives „Gedächtnis“, in dem der Verlauf der „Energiefläche“ in einem größeren, nichtlokalen Bereich gespeichert ist. Ein

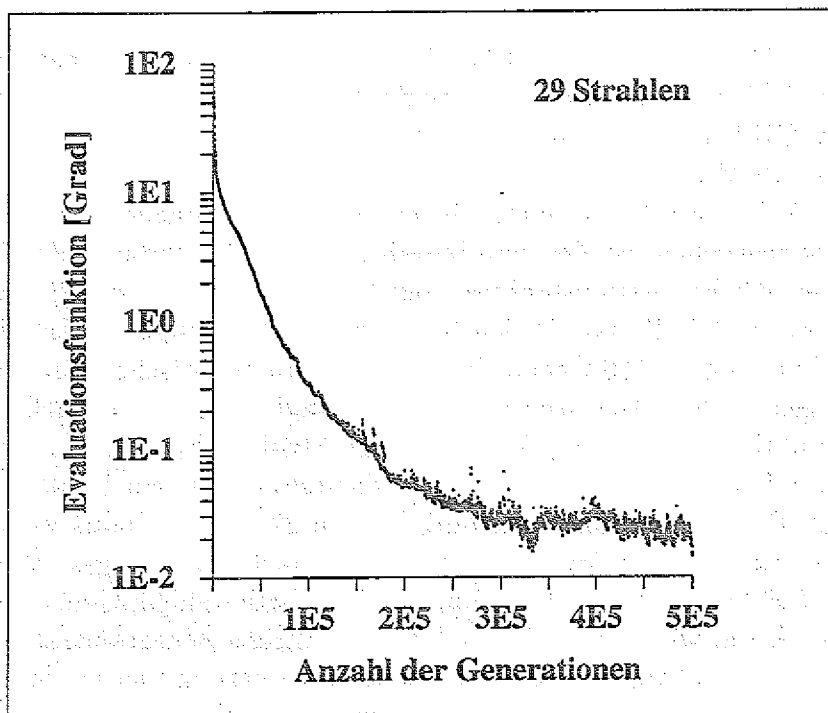
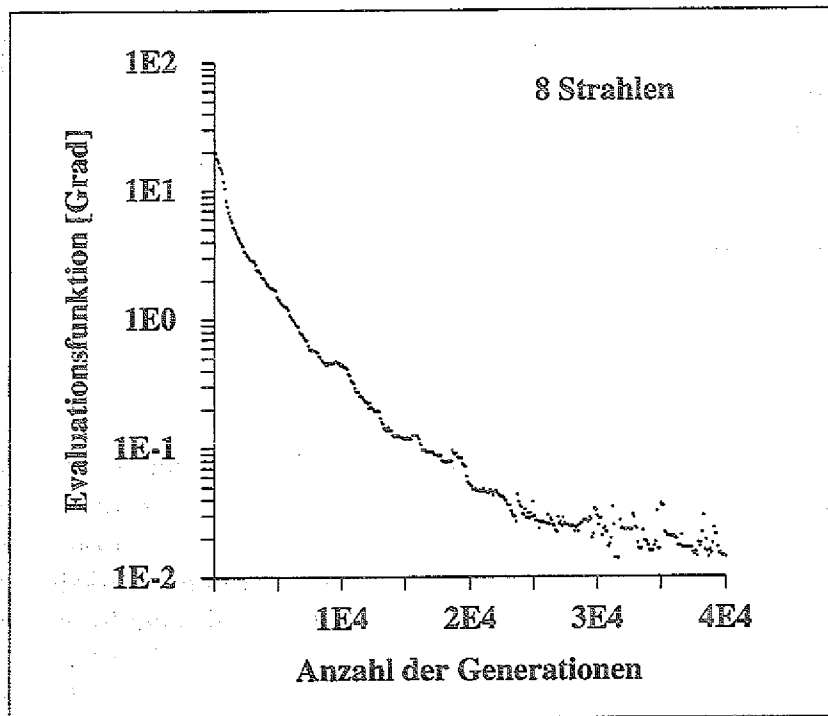


Abbildung 8.7: Konvergenzdarstellungen, die typische Programmläufe des genetischen Algorithmus zeigen. Oben: 8-Strahl Modus, unten: 29-Strahl Modus. Der zum jeweils „besten“ Individuum der Population gehörende Wert der Evaluationsfunktion ist logarithmisch über der Anzahl der Generationen aufgetragen.

Verlust an Vielfalt wird hierbei durch den Mutationsoperator verhindert. Er erzeugt stets Individuen, die außerhalb des jeweiligen lokalen Minimums liegen. Bei zu hoher Mutationsrate hingegen wird das kollektive Gedächtnis durch zufällige Änderungen eines Großteils der Individuen laufend zerstört. Die daraus resultierende Dominanz der globalen Erkundung des Lösungsraumes verhindert jegliche lokale Optimierung und erzeugt chaotische Sprünge des aktuellen Bestwertes der Evaluationsfunktion. Die Mutationsrate muß deshalb so angepaßt werden, daß die globale Exploration und die lokale Optimierung in einem ausgewogenen Verhältnis zueinander stehen.

8.4 Rekonstruktion der Austrittswellenfunktion aus experimentellen Abbildungen von Ni_4Mo

Die Eignung der Algorithmen, die Austrittswellenfunktion aus nur einigen wenigen Eingabebildern zu rekonstruieren, wird in diesem Abschnitt anhand der in Kapitel 7 gezeigten experimentellen Defokus-Serie untersucht. Um die Rekonstruktion möglichst schwierig zu gestalten, werden anstatt der gesamten Serie nur die partiell symmetriekorrigierten Bilder 3 und 4 aus Abbildung 7.6 als Eingabe benutzt. Die vorherige Symmetriekorrektur ermöglicht zum einen den rechenzeitsparenden Einsatz des 8-Strahl Modus der Algorithmen. Zum anderen ist die aufgrund der aufgeprägten Symmetriebedingung erhaltene vierzählige Lösung besser mit simulierten Wellenfunktionen, die perfekte Symmetrie aufweisen, vergleichbar.

Amplitude und Phase der mittels des Simulated-Annealing-Algorithmus rekonstruierten Wellenfunktion sind in Abbildung 8.8a dargestellt. Die entsprechende Lösung, die durch Einsatz des genetischen Algorithmus ermittelt wurde, ist in Abbildung 8.8b zu sehen. Die von den beiden verschiedenen Algorithmus-Typen ermittelten Wellenfunktionen sind optisch identisch. Mit keinem der beiden Algorithmen war es möglich, Details der Wellenfunktion, die kleiner als 0.18 nm sind, aufzulösen. Aus diesem Grund ist die Auflösung für Amplitude und Phase in Abbildung 8.8 auf diesen Wert beschränkt. Beide Algorithmen ermittelten für das Minimum der Evaluationsfunktion einen Wert von 3.8° . Die Ursache für diesen relativ hohen Wert wird im nächsten Abschnitt diskutiert.

Die rekonstruierte Amplitude, die in den Abbildungen 8.8a und 8.8b jeweils oben gezeigt wird, ist eng mit der Struktur der kubisch flächenzentrierten Matrix verknüpft, die der Ni_4Mo Struktur zugrundeliegt. Orte minimaler Amplitude liegen an den projizierten Atompositionen der Matrix, Orte maximaler Amplitude liegen entsprechend an den dazwischenliegenden Tunnelpositionen. Schwächer ausgeprägte Kontrastdetails in den Amplitudenbildern deuten die Abfolge von vier Nickelatomen und einem Molybdänatom entlang den $\langle 100 \rangle$ Kristallrichtungen der kfz Matrix an. Die in die Nickelmatrix eingebettete Überstruktur der Molybdänatome ist in der Phase der Wellenfunktion, die in der unteren Reihe der Abbildungen 8.8a und 8.8b dargestellt ist, klar zu erkennen. Die hell dargestellten, kreisförmig umrandeten Maxima der Phase liegen an projizierten Molybdänpositionen.

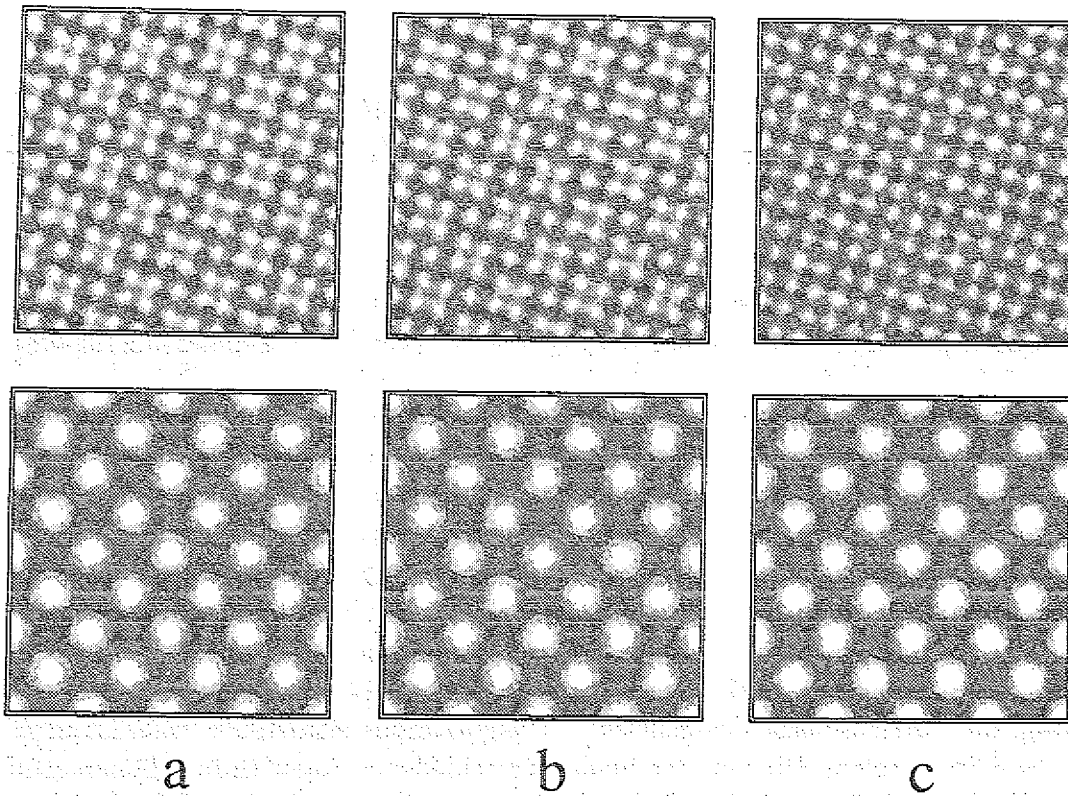


Abbildung 8.8: (a) Amplitude (oben) und Phase (unten) der Austrittswellenfunktion, die mittels des Simulated-Annealing-Algorithmus aus den Bildern 3 und 4 der experimentellen Defokus-Serie in Abbildung 7.6 ermittelt wurden. (b) Amplitude und Phase, die aus denselben Bildern mittels des genetischen Algorithmus bestimmt wurden. (c) Wellenfunktion, die durch numerische Simulation des dynamischen Elektronenbeugungsprozesses berechnet wurde. Die Simulation basiert auf einer Objektdicke von 14 nm.

Um die Verlässlichkeit der Algorithmen im praktischen Einsatz weiter zu untersuchen, wurden die in den Abbildungen 8.8a und 8.8b dargestellten rekonstruierten Wellenfunktionen mit solchen verglichen, die mittels numerischer Simulation der Elektronenbeugung berechnet wurden. Mit Hilfe des EMS Pakets [46] wurden Wellenfunktionen für eine Einstrahlrichtung entlang der [001] Zonenachse im Objektdickenbereich von 0.5 bis 30 nm in Schritten von 0.5 nm berechnet. Die für die jeweiligen Objektdicken berechneten Wellenfunktionen unterscheiden sich drastisch innerhalb relativ geringer Dickenänderungen von 1 bis 2 nm. Amplitude und Phase der für eine Objektdicke von 14 nm simulierten Wellenfunktion sind in Abbildung 8.8c dargestellt. Die dominanten Kontrastmerkmale der simulierten Amplitude sind ähnlich wie die der rekonstruierten Amplitude. Die schwache Andeutung der vier-zu-eins Abfolge von Nickel- und Molybdänatomen, wie sie in der Rekonstruktion der Amplitude zu Tage tritt, findet sich jedoch nicht in der Simulation der Amplitude. Andererseits entspricht die rekonstruierte Phase perfekt dem entsprechenden Simulationsergebnis. Außer bei einer Objektdicke um 14 nm tritt in der Simulation keine andere Wellenfunktion auf, die auch nur annähernd so gut mit der rekonstruierten Wellenfunktion übereinstimmt.

8.5 Diskussion

Im Rahmen dieser Arbeit wird die Anwendung der beiden stochastischen Algorithmen zur Rekonstruktion der Wellenfunktion auf periodische Bildmotive beschränkt. Es gibt keinen prinzipiellen Grund für diese Beschränkung, da die Struktur der verwendeten Algorithmen vollkommen unabhängig von der Art der Eingabeobjekte ist. Der Grund für die Behandlung periodischer Motive beruht auf der Tatsache, daß sich der Informationsgehalt solcher Motive im Fourierraum durch relativ wenige Amplituden- und Phasenwerte beschreiben läßt. In der Terminologie des Simulated-Annealing-Algorithmus bedeutet dies, daß die Anzahl der zu behandelnden Freiheitsgrade auf eine Größenordnung kleiner 100 beschränkt ist. Im Gegensatz zu periodischen Motiven ist der Informationsgehalt nicht-periodischer Bilder über größere Gebiete des Fourierraums kontinuierlich verteilt. In Abhängigkeit vom Bildinhalt ist die Anzahl der beteiligten Freiheitsgrade meist weitaus größer als im periodischen Fall. Der Rechenaufwand für die Rekonstruktion der Austrittswellenfunktion steigt jedoch exzessiv mit der Anzahl der beteiligten Fourierkomponenten an.

Für den hohen Rechenaufwand, der für die Rekonstruktion nicht-periodischer Bildinhalte mittels der gegenwärtigen Implementation der Algorithmen zu erwarten ist, sind zwei Gründe verantwortlich. Zum einen steigt die Anzahl der erforderlichen Aufrufe der Evaluationsfunktion überproportional mit der Anzahl der an der Rekonstruktion beteiligten Freiheitsgrade an. Dies wird bei Betrachtung der zum genetischen Algorithmus gehörenden Konvergenzdarstellungen in Abbildung 8.7 deutlich. Die Anzahl der Generationen ist bei etwa gleichem Wert der Evaluationsfunktion im 29-Strahl Modus etwa zehnmal höher als im 8-Strahl Modus. Ein linearer Anstieg der erforderlichen Aufrufe der Evaluationsfunktion würde jedoch nur einer vierfach größeren Anzahl von Generationen entsprechen. Beim Simulated-Annealing-

Algorithmus wurde bei steigender Strahlzahl ein ähnlich starker Anstieg der erforderlichen Konfigurationsänderungen festgestellt. Die genaue Form des Anstiegs hängt vom Bildmotiv selbst ab. Zwei Extreme sind möglich: Im Falle vornehmlich linearen Bildkontrasts, der annähernd dann zustandekommt, wenn die Intensität des Nullstrahls viel größer als die der abgebeugten Strahlen ist, ist mit einem linearen Anstieg der Rechenzeit in Abhängigkeit von der Strahlzahl zu rechnen. Andererseits muß bei starker Kopplung der abgebeugten Strahlen untereinander aufgrund des kombinatorischen Charakters des zugrundeliegenden Problems im schlimmsten Fall mit einem exponentiellen Anstieg der Rechenzeit gerechnet werden.

Zahlreiche Anwendungen des Simulated-Annealing-Algorithmus sowie des genetischen Algorithmus auf das gleiche Rekonstruktionsproblem zeigten, daß die Sucheffizienz der beiden Algorithmen etwa gleich ist. Beide Algorithmen sind in ihrem Kern sehr einfach zu implementieren, doch gibt es einen Unterschied in der Nutzerfreundlichkeit.

Für das *simulated annealing* müssen zwei Parameter, nämlich die Anfangstemperatur sowie die Konstante, die den Temperaturabfall beschreibt, richtig justiert werden. Das mögliche Einfrieren des Algorithmus in einem lokalen Minimum der Evaluationsfunktion kann sehr einfach nach einem Neustart des Programms verhindert werden, wobei entweder eine höhere Anfangstemperatur oder ein moderaterer Temperaturabfall als zuvor gewählt werden. Die richtige Bandbreite der Fluktuationen in der Evaluationsfunktion wird dann durch den extern vorgegebenen „Temperaturfahrplan“ während des Programmlaufs laufend angepaßt. Bei der Anwendung des genetischen Algorithmus ist die Wahl einer effizienten Populationsgröße und einer geeigneten Mutationsrate erforderlich. Diese beiden Parameter bestimmen die Bandbreite der Fluktuationen für die jeweiligen Bestwerte der Evaluationsfunktion. Die Fluktuationsbandbreite bleibt während des Programmlaufes im Gegensatz zum *simulated annealing* unverändert. Die Wahl einer geeigneten Bandbreite ist von der Rauigkeit der „Energiefläche“ abhängig. Die nötige Feinjustage von Populationsgröße und Mutationsrate ist deshalb schwieriger als das Festlegen des entsprechenden „Temperaturfahrplans“ beim Simulated-Annealing-Algorithmus.

Eine sehr interessante Frage ist die nach der minimalen Anzahl der für die Rekonstruktion der Austrittswellenfunktion benötigten Bilder. Wie in Abschnitt 8.3.1 berichtet wurde, ist die Rekonstruktion der Wellenfunktion aus dem in Abbildung 8.4 gezeigten Bild 2 alleine möglich. Dies erscheint auf den ersten Blick paradox. Angenommen, das Bild sei auf einem diskreten Gitter digitalisiert, dann kann man für jeden Abtastpunkt nur einen Zahlenwert, nämlich die Bildintensität, messen. Für die Austrittswellenfunktion müssen jedoch *zwei* Werte, nämlich Amplitude und Phase bekannt sein.

Das oben beschriebene „Paradoxon“ ist einfach aufzulösen. Dies sei anhand des idealisierten Beispiels eines eindimensionalen Bildes, das unter kohärenter Beleuchtung aufgenommen wurde und bei dem Dämpfungseffekte keine Rolle spielen, erläutert. Die Wellenfunktion setze sich aus N Fourierkomponenten zusammen. Aufgrund der Autokorrelation des Fourierspektrums (siehe Anhang A) enthält das Bild dann $2N - 1$ Fourierkomponenten. Da die Fouriertransformierte des Bildes der Friedelsymmetrie unterliegt, können also $(2N - 2)/2 + 1 = N$ Amplituden- und

Phasenwerte gemessen werden. Dies entspricht genau der Anzahl der Unbekannten. Diese Erklärung hat ein Analogon im Realraum: Wegen der Verdoppelung des Fourierspektrums durch die Autokorrelation, muß das Bild im Realraum wegen des Whittaker-Shannon Theorems [83] mit der doppelten Abtastrate wie die Wellenfunktion erfaßt werden. Legt man das Abtastfeld des Bildes zugrunde, so ist die Wellenfunktion auf diesem Feld unnötig dicht erfaßt, es würde jeder zweite Punkt ausreichen um die volle Information zu speichern. Zum Beispiel genügt es, für jeden geraden Bildpunkt die Phase der Wellenfunktion zu ermitteln und für jeden ungeraden die Amplitude. Auch im Realraum steht also für jede Unbekannte im Mittel ein Meßwert zur Verfügung.

Die erfolgreiche Anwendung der stochastischen Algorithmen läßt sogar vermuten, daß die in Bild 2 der Abbildung 8.4 enthaltene Information teilweise redundant ist. Die im Bildmittelwert enthaltene Information scheint überflüssig zu sein, da der Bildmittelwert nicht in den Kreuzkovarianzkoeffizienten (Gl. (7.4)) eingeht und somit in der Evaluationsfunktion (Gl. (8.1)) keine Rolle spielt. Dies trifft auch auf alle Rekonstruktionen aus zwei oder mehr Bildern zu. Die hier gewählte Methode des Bildvergleichs läßt den Mittelwert der Bildintensität vollkommen außer Acht. Trotzdem wurde nie ein Scheitern der Rekonstruktion festgestellt.

Weicht man von der obigen Idealisierung vollständiger Kohärenz ab, ist die Frage nach der Anzahl der benötigten Eingabebilder nicht mehr generell zu beantworten. Die Antwort hängt dann vom Raumfrequenzspektrum der zu untersuchenden Struktur sowie von den gewählten Abbildungsbedingungen ab. Eine Fourierkomponente der Wellenfunktion kann sicherlich nicht rekonstruiert werden, wenn sie keinen Einfluß auf den Bildkontrast einer gesamten Defokus-Serie hat. Eine solche Situation ist beispielsweise dadurch gegeben, daß die zugehörige Raumfrequenz jenseits des Informationslimits liegt und die Dämpfung aufgrund der partiell kohärenten Beleuchtung vollständig ist. Abbildungsbedingungen, die nahe an dieser Situation sind, werden an späterer Stelle quantitativ diskutiert. Es ist ebenso möglich, daß Übertragungslücken sowohl im linearen als auch im nichtlinearen Kontrasttransfer aus den Abbildungen Information, die zur Rekonstruktion der Wellenfunktion nötig ist, ausschließen. Dies ist umso wahrscheinlicher, je geringer die Anzahl der Eingabebilder ist. Das Problem kann jedoch prinzipiell immer durch eine höhere Zahl von Eingabebildern gelöst werden. Eine höhere Zahl von Eingabebildern als die unbedingt benötigte ist generell empfehlenswert, da der Einfluß experimentellen Rauschens durch die Einbeziehung redundanter Information reduziert wird.

Speziell eine theoretisch mögliche Rekonstruktion der Wellenfunktion aus einem einzeltem Eingabebild ist in Gegenwart von experimentellem Rauschen wahrscheinlich nicht mehr möglich. Bei der Rekonstruktion der Wellenfunktion aus dem Bild 2 in Abbildung 8.4 wurde nämlich festgestellt, daß bei Werten der Evaluationsfunktion von 0.1° noch eine andere Wellenfunktion existiert, die sich gänzlich von der richtigen Lösung unterscheidet. Diese andere Wellenfunktion entspricht einem lokalen Minimum der Evaluationsfunktion, das zwar im Lösungsraum weit entfernt vom globalen Minimum ist, dessen Funktionswert jedoch nur etwas größer als der minimale Wert von 0° ist. Nun ist es wahrscheinlich, daß das Rauschen einen Beitrag zur Evaluationsfunktion stellt, der einer Abweichung von mehr als 0.1° vom idealen Bild

entspricht. Somit würde das experimentelle Rauschen im Bild zufällig entscheiden, welche von den beiden „Lösungen“ gefunden wird. Die Anzahl der lokalen Minima der Evaluationsfunktion kann durch Erhöhung der Anzahl der Eingabebilder enorm reduziert werden. Dies wird dadurch deutlich, daß eine Rekonstruktion aus dem in Abbildung 8.4 gezeigten Bild 2 alleine etwa zehnmal mehr Rechenzeit erfordert als die Rekonstruktion aus den beiden dort gezeigten Bildern. Dies ist umso bemerkenswerter, wenn man beachtet, daß ersten Fall die Berechnung der Evaluationsfunktion aufgrund des einzelnen Eingabebildes nur halb solange wie bei der Rekonstruktion aus beiden Bildern dauert. Bei einer Rekonstruktion aus zwei Bildern wird also die Anzahl der lokalen Minima der Evaluationsfunktion im Vergleich zur Rekonstruktion aus einem Bild so stark vermindert, daß nur mehr etwa ein zwanzigstel des ursprünglichen Suchaufwandes erforderlich ist.

Es gibt generell die Möglichkeit, daß zwei unterschiedliche Austrittswellenfunktionen eine identische Defokus-Serie erzeugen können. Detaillierte analytische Betrachtungen dieses Problems gibt es nach Kenntnis des Verfassers nur für den eindimensionalen Fall unter der Annahme kohärenter Einstrahlung [112]. Diese trotz der Vereinfachungen sehr anspruchsvollen mathematischen Manipulationen zur Analyse dieses Problems übersteigen den Rahmen der vorliegenden Arbeit. Mit Hilfe der hier besprochenen stochastischen Algorithmen ist es jedoch möglich, durch mehrfache Programmläufe die Existenz von mehrdeutigen Lösungen zu untersuchen. Die Abläufe des Simulated-Annealing-Algorithmus sowie des genetischen Algorithmus werden zu jedem Zeitpunkt durch Zufallszahlen bestimmt, die durch einen Zufallszahlengenerator erzeugt werden. Da in einen Programmlauf etwa 10^5 Zufallszahlen eingehen, erfolgt die Suche im Raum der möglichen Lösungen niemals entlang desselben Pfades, wenn unterschiedliche Startwerte des Zufallszahlengenerators benutzt werden. Wenn beispielsweise zehn unterschiedliche Programmläufe identische Lösungen liefern, wie das bei den hier gezeigten Beispielen der Fall ist, so ist die Existenz zusätzlicher, nicht entdeckter Lösungen sehr unwahrscheinlich.

Wie in den Tabellen 8.1 und 8.2 zu sehen ist, sind die von den Algorithmen ermittelten Ergebnisse von fast analytischer Genauigkeit. Für Strahlen niedriger Ordnung sind die Fehler der Amplituden und der Phasen kaum noch numerisch auflösbar. Die Lösung für Strahlen höherer Ordnung, wie für die {330}- und die {400} Strahlen, weist jedoch nicht vernachlässigbare Fehler auf. Diesen durch die dynamische Elektronenbeugung nur schwach angeregten Strahlen entsprechen Raumfrequenzen, die jenseits der Punktauflösung des Mikroskops liegen. Aufgrund der Dämpfungseffekte, die durch die partielle Kohärenz des Elektronenstrahls hervorgerufen werden, ist der Beitrag dieser Strahlen zum Bildkontrast äußerst gering. Die quantitative Behandlung solch schwacher Kontrastdetails stellt hohe Anforderungen an das Konvergenzverhalten der Algorithmen.

Diese Anforderungen an das Konvergenzverhalten wurden anhand der Testbilder aus Abbildung 8.4 quantitativ untersucht. In Tabelle 8.3 sind diejenigen Abweichungen der Evaluationsfunktion vom minimalen Wert 0° vermerkt, die durch einen relativen Fehler der Strahlamplituden von jeweils 10 % oder durch einen absoluten Fehler der Phasen von jeweils 5° hervorgerufen werden. Während bei den Strahlen niedriger Ordnung relative Amplitudenabweichungen von 10 % deutliche

(hkl)	$A + 10\%$	$A - 10\%$	$\varphi + 5^\circ$	$\varphi - 5^\circ$
(000)	4.548°	5.184°	5.545°	5.211°
(110)	4.979°	5.138°	4.305°	4.273°
(200)	4.332°	4.652°	4.057°	4.495°
(220)	1.240°	1.263°	1.953°	1.936°
(310)	1.401°	1.378°	0.888°	0.890°
(130)	2.778°	2.810°	2.114°	2.090°
(400)	0.046°	0.042°	0.034°	0.034°
(330)	0.005°	0.001°	0.002°	0.006°

Tabelle 8.3: Abweichungen der Evaluationsfunktion vom minimalen Wert 0° , die durch einen relativen Fehler der Amplitude von 10 % oder einen absoluten Fehler der Phase von 5° hervorgerufen werden. Die Abweichungen sind jeweils für einen gemeinsamen Fehler aller vier Strahlen berechnet, die zu einer symmetrieverwandten Gruppe (hkl) gehören.

Änderungen der Evaluationsfunktion bewirken, haben 10 %-ige Amplitudenabweichungen der $\{330\}$ - und $\{400\}$ Strahlen eine 2 bis 3 Größenordnungen geringere Auswirkung auf den Wert der Evaluationsfunktion. Ein ähnliches Verhalten besteht für die Auswirkungen von absoluten Phasenabweichungen von 5° um die richtige Lösung. Der Einfluß der $\{330\}$ - und $\{400\}$ Strahlen auf den Bildkontrast ist auch hierbei um Größenordnungen geringer als der von Strahlen niedriger Ordnung. Für einen bestimmten Strahl kann eine Lösung mit einem Amplitudenfehler, der kleiner als 10 % ist, oder eine Lösung mit einem Phasenfehler, der kleiner als 5° ist, nur dann erwartet werden, wenn die in Tabelle 8.3 aufgelisteten Werte jeweils kleiner als 10^{-2} Grad sind. Dieser Wert entspricht dem numerischen Genauigkeitslimit der in den Programmen verwendeten Zahlendarstellung. Es ist mehr als wahrscheinlich, daß schwache Kontrastbeiträge wie die der $\{330\}$ - und $\{400\}$ Strahlen in experimentellen Bildern aufgrund von Rauschen nicht mehr meßbar sind. Deswegen wäre eine Steigerung der in den Programmen benutzten Zahlendarstellungsgenauigkeit ohne praktischen Nutzen.

Der Endwert der Evaluationsfunktion, der bei der Rekonstruktion der in Abbildung 8.8 dargestellten experimentellen Wellenfunktion erzielt wurde, beträgt 3.8° . Dieser Wert, der zwei Größenordnungen über typischen Werten liegt, die bei der Rekonstruktion simulierter Bilder erreicht wurden, bringt zum Ausdruck, daß schwache Kontrastdetails der experimentellen Bilder nicht von ein und derselben Wellenfunktion erzeugt werden können. Neben experimentellem Rauschen sind für die widersprüchliche Information hauptsächlich diejenigen Fourierkomponenten des Bildes 4 in Abbildung 7.6 verantwortlich, die bei der Symmetriekorrektur aufgrund der in Gl. (7.22) aufgestellten Bedingungen nicht berücksichtigt werden konnten. Es ist möglich, daß dies die Ursache für die leichte Diskrepanz zwischen rekonstruierter Amplitude und der durch Simulation der dynamischen Elektronenbeugung errechneten Wellenamplitude ist.

Ein allen Rekonstruktionstechniken gemeinsames Problem ist die Tatsache, daß die experimentellen Abbildungsparameter genau bekannt sein müssen. Für die Bestimmung von Semikonvergenzwinkel und Defokusstreuung existieren ausreichend genaue Meßmethoden [58, 64]. Ebenso wurden in den letzten Jahren Methoden zur Präzisionsmessung der sphärischen Aberrationskonstante entwickelt [122, 123]. Ein noch ungelöstes Problem, von dem die Verlässlichkeit der mittels einer Rekonstruktionsmethode gewonnenen Daten empfindlich abhängt, ist die Ermittlung der zu den Eingabebildern gehörenden *absoluten* Defokuswerte. Die in dieser Arbeit für die Rekonstruktion der Austrittswellenfunktion aus experimentellen Bildern benutzten Defokuswerte wurden durch Vergleich mit simulierten Bildern bestimmt (siehe Kapitel 7).

Mit der hier diskutierten Rekonstruktionsmethode wurde der Versuch durchgeführt, zusammen mit der Wellenfunktion auch die absoluten Defokuswerte simulierter Eingabebilder zu bestimmen. Es zeigte sich, daß dies nicht möglich ist. Der Grund hierfür liegt in der Tatsache, daß die Wellenpropagation im Elektronenmikroskop in zwei Stufen abläuft. Zum einen wird die Wellenpropagation innerhalb des Objekts mittels des Propagatorterms (siehe Gl. (4.49)) beschrieben, zum anderen wird die Propagation der Welle von der Objekzebene zur Bildebene durch den Defokusterm der Aberrationsfunktion dargestellt (siehe Gl. (5.8)). Da die beiden Anteile der Wellenpropagation im Bild nicht mehr zu trennen sind, ist es nicht möglich festzustellen, welcher Anteil der Propagation auf die Defokussierung zurückgeht. Da die Bilder einer Defokus-Serie jedoch bezüglich der Summe der beiden Propagationsanteile konsistent sein müssen, sollte es mittels der diskutierten Methode möglich sein, *relative* Defokusabstände innerhalb einer Bildserie zu bestimmen.

8.6 Optimierung der numerischen Effizienz

In der Diskussion zu den vorangehenden Abschnitten dieses Kapitels wurde festgestellt, daß die *Sucheffizienz* der beiden zur Rekonstruktion der Austrittswellenfunktion benutzten Algorithmen etwa gleich ist. Abgesehen vom Unterschied in der Nutzerfreundlichkeit benötigen der Simulated-Annealing-Algorithmus und der genetische Algorithmus etwa gleich viele Aufrufe der Evaluationsfunktion. Die Berechnung der Evaluationsfunktion nimmt den Großteil des eingesetzten Rechenaufwands in Anspruch. Wie bereits diskutiert wurde, kann bei der Behandlung komplexer Strukturen, bei denen eine große Anzahl von Fourierkomponenten eine Rolle spielt, sowie bei starker nichtlinearer Kopplung der beitragenden Strahlen die Anzahl der benötigten Aufrufe der Evaluationsfunktion drastisch zunehmen. Umso wichtiger ist es in solchen Fällen, die pro Evaluation eingesetzte Rechenzeit zu minimieren.

Die im vorangehenden eingesetzte Routine zur Berechnung der Evaluationsfunktion war für beide Algorithmustypen dieselbe. Diese Routine besteht im wesentlichen aus zwei Teilen. Zum einen werden aus der von den Algorithmen „vorgeschlagenen“ Testwellenfunktion die zu den jeweiligen Defokuswerten gehörenden Bilder berechnet. Anschließend erfolgt der Vergleich der so berechneten Bilder mit den experimentellen Eingabebildern. Der im zweiten Schritt erfolgende Bildvergleich nimmt den

wesentlich kleineren Teil der Rechenzeit in Anspruch, da die Anzahl der zu vergleichenden Fourierkomponenten des Bildes proportional zur Anzahl N der Fourierkomponenten der Wellenfunktion ist. Die im ersten Schritt erfolgende Bildberechnung ist wesentlich aufwendiger, da sie die Berechnung von N^2 Interferenztermen, die zu jedem Bild beitragen, erfordert.

Vom numerischen Standpunkt ergibt sich hier ein weiterer nicht zu unterschätzender Vorteil für den Simulated-Annealing-Algorithmus. Während beim genetischen Algorithmus pro Iterationsschritt durch den Paarungsoperator im Mittel die Hälfte der Fourierkoeffizienten der Versuchswellenfunktion eine Änderung erfährt, wird beim Simulated-Annealing-Algorithmus pro Konfigurationswechsel jeweils nur eine Fourierkomponente ausgetauscht. In letzterem Fall muß nicht das ganze Bild entsprechend N^2 Interferenztermen neu berechnet werden, sondern es müssen nur die $2N$ Interferenzterme, die sich durch die Modifikation einer einzigen Fourierkomponente der Wellenfunktion ändern, auf den neuesten Stand gebracht werden. Diese im Englischen „*differential updating*“ genannte Programmieretechnik, die beim Einsatz des Simulated-Annealing-Algorithmus genutzt werden kann, reduziert den Rechenaufwand pro Evaluation drastisch. War bei der früher benutzten Version der Evaluationsroutine der Rechenaufwand proportional zu N^2 , so ist bei der für den Simulated-Annealing-Algorithmus speziell angepaßten Version der Rechenaufwand nur mehr proportional zu N .

Die Fourierdarstellung der Bildintensität $I_{n+1}(\vec{k})$ der $n+1$ -ten Konfiguration ergibt sich durch die Berechnung der Differenz $\Delta I(\vec{k})$ zur Fourierdarstellung der n -ten Konfiguration $I_n(\vec{k})$:

$$I_{n+1}(\vec{k}) = I_n(\vec{k}) + \Delta I(\vec{k}). \quad (8.4)$$

Wird die zu einer bestimmten Raumfrequenz $\vec{g}$ gehörende Fourierkomponente $F(\vec{g})$ beim Konfigurationswechsel durch den Algorithmus geändert, so lautet der Differenzterm für $\vec{k} \neq 0$

$$\Delta I(\vec{k}) = \Delta F(\vec{g}) F^*(\vec{g} - \vec{k}) T(\vec{g}, \vec{g} - \vec{k}, Z) + F(\vec{g} + \vec{k}) \Delta F^*(\vec{g}) T(\vec{g} + \vec{k}, \vec{g}, Z), \quad (8.5)$$

wobei

$$\Delta F(\vec{g}) = F_{n+1}(\vec{g}) - F_n(\vec{g}). \quad (8.6)$$

Bei einem Lauf des Simulated-Annealing-Algorithmus werden typischerweise 10^5 bis 10^6 Konfigurationswechsel durchgeführt. Durch das ständige Bilden der Differenz würde im Zusammenspiel mit Rundungsfehlern die Testwellenfunktion während des Programmlaufs mehr und mehr verfälscht. Deshalb kann die „*differential updating*“ Technik nicht in Reinform angewandt werden. Von Zeit zu Zeit müssen kulminierende Rundungsfehler durch „Einschieben“ einer vollen Bildberechnung nach Gleichung (7.2) wieder eliminiert werden. Es zeigte sich jedoch, daß die volle Bildberechnung frühestens nach 3000 – 5000 „*updating*“-Schritten erforderlich ist, sodaß dieser zusätzliche numerische Aufwand kaum merklich ist.

Bei den im nächsten Abschnitt besprochenen Rekonstruktionen an Bildern von $\text{YBa}_2\text{Cu}_3\text{O}_7$, bei denen 53 Fourierkomponenten der Wellenfunktion berücksichtigt

werden, macht sich die Anwendung der hier besprochenen Modifikation in einer Geschwindigkeitssteigerung von etwa einem Faktor 10 gegenüber der früheren Version bemerkbar. Die Rekonstruktion der aus 53 Strahlen bestehenden Wellenfunktion mittels des Simulated-Annealing-Algorithmus bis zu einem Wert der Evaluationsfunktion von $4 \cdot 10^{-2}$ Grad benötigt auf einer DEC 3500 ALPHA Rechenanlage nur etwa fünf Minuten.

8.7 Rekonstruktion der Austrittswellenfunktion aus Abbildungen von $\text{YBa}_2\text{Cu}_3\text{O}_7$

In den Abbildungen 8.9 und 8.10 sind zwei Ausschnitte aus Hochauflösungsaufnahmen von $\text{YBa}_2\text{Cu}_3\text{O}_7$ dargestellt, die entweder entlang der [100] oder der [010] Zonenachse aufgenommen wurden. Diese beiden Projektionsrichtungen unterscheiden sich nur durch die Ausrichtung der CuO Ketten, die entlang der b-Achse der Einheitszelle verlaufen. Im Falle der [100] Zonenachse verlaufen die CuO Ketten senkrecht zum einfallenden Strahl, im Falle der [010] Zonenachse jedoch parallel zum einfallenden Strahl (Siehe Abb. 2.5). Da die Sauerstoffatome aufgrund ihrer niedrigen Kernladungszahl einfallende Elektronen nur sehr schwach streuen, ist in Hochauflösungsaufnahmen eine Unterscheidung der beiden Projektionsrichtungen nur bei Objektdicken von mehr als 8 nm möglich [35, 36]. Da es im folgenden auf eine Unterscheidung der Zonenachsen nicht ankommt, wird deshalb der Einfachheit halber von der [100] Zonenachse gesprochen.

Die in den Abbildungen 8.9 und 8.10 gezeigten Bilder bestehen aus 512×512 Bildpunkten und entsprechen der in Kapitel 7 erläuterten ersten Digitalisierungsstufe der quantitativen Bildvergleichsmethode. Die Bildausschnitte gehören zu zwei Aufnahmen, die an derselben Objektstelle mit unterschiedlichem Defokus aufgenommen wurden. Der Durchmesser der gewählten Objektiveblende entspricht 11 nm^{-1} , sodaß jeweils 53 Strahlen zu den Abbildungen beitragen (Abb. 3.8). Die zur Rekonstruktion der Austrittswellenfunktion ausgewählten Objektgebiete sind in den Abbildungen 8.9 und 8.10 durch einen Rahmen gekennzeichnet. Da keine nicht-periodische Stelle aufzufinden war, die das Auffinden exakt derselben Probenstelle auf den beiden Aufnahmen ermöglicht, sind die „wahren“ Orte der ausgewählten Bereiche möglicherweise um einige Einheitszellen verschoben. Weil jedoch keine kurzreichweitige systematische Änderung des Bildkontrasts innerhalb der aus 512×512 Bildpunkten bestehenden Bereiche zu beobachten ist, dürfte eine Verschiebung um einige Einheitszellen bei der Rekonstruktion kaum ins Gewicht fallen.

Um die Defokuswerte zu ermitteln, unter denen die Abbildungen 8.9 und 8.10 aufgenommen wurden, kommt die in Kapitel 7 vorgestellte Bildvergleichsmethode zur Anwendung. Die Bilder 1 und 2 der Abbildung 8.11 zeigen die eingerahmten Objektgebiete aus den Abbildungen 8.9 und 8.10 nach der Fourierraum-Mittelung. Diese Bearbeitungsstufe entspricht den in der mittleren Spalte von Abbildung 7.6 gezeigten Bildern von Ni_4Mo . Bild 3 der Abbildung 8.11 ist dasjenige simulierte Bild, das bei Variation von Defokus und Objektdicke den höchsten Wert der Kreuzkovarianz zum experimentellen Bild 1 aufweist. Entsprechend ist Bild 4 der Abbil-

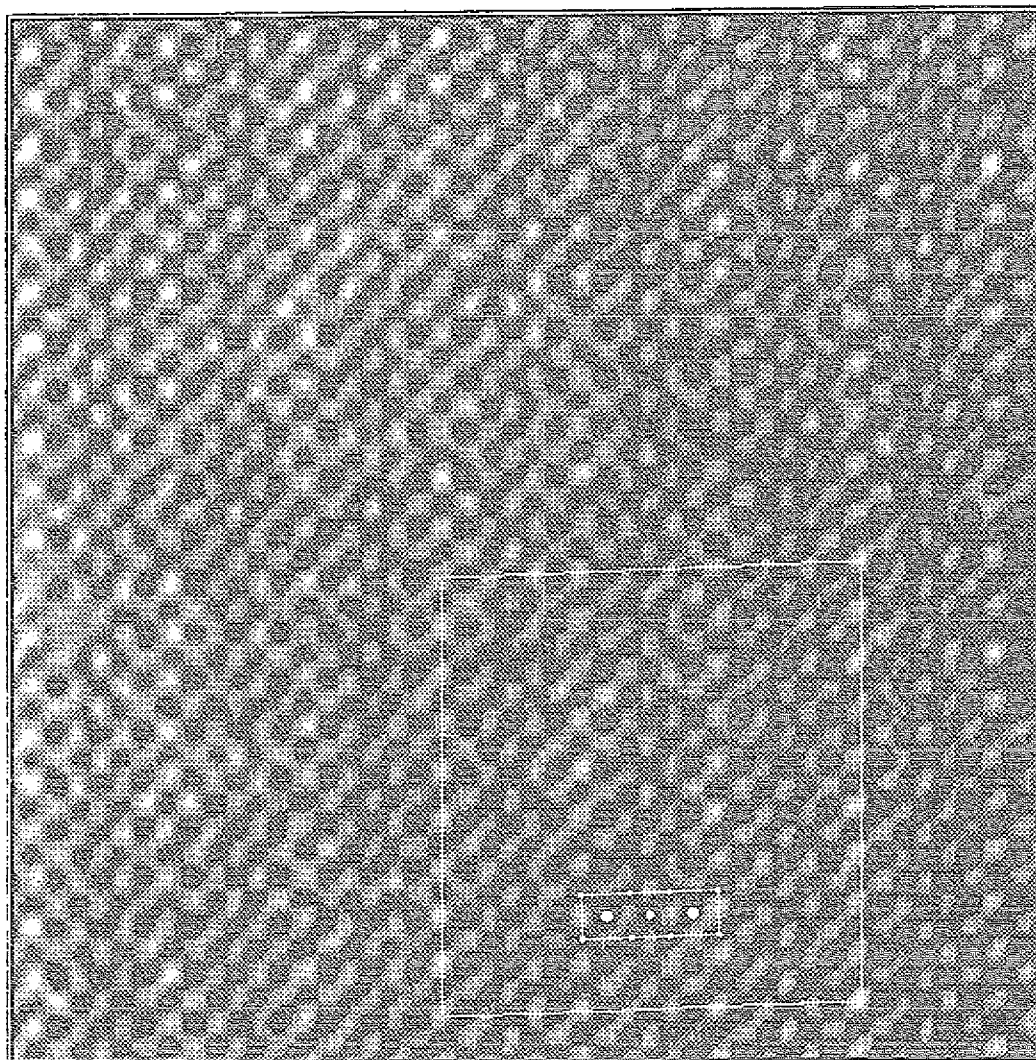


Abbildung 8.9: Hochauflösende Abbildung von $\text{YBa}_2\text{Cu}_3\text{O}_7$ entlang der $[100]$ Zonenachse. Die Abbildung besteht aus 512×512 Bildpunkten. Das zur Rekonstruktion der Wellenfunktion ausgewählte Objektgebiet ist durch einen Rahmen gekennzeichnet. Eine Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ ist innerhalb des Rahmens eingezeichnet. Kleine Kreise bezeichnen Cu-Positionen, der mittlere die Y-Position, und große Kreise die Ba-Positionen der Einheitszelle. Die Kantenlänge des eingezeichneten Rahmens beträgt 3.5 nm.

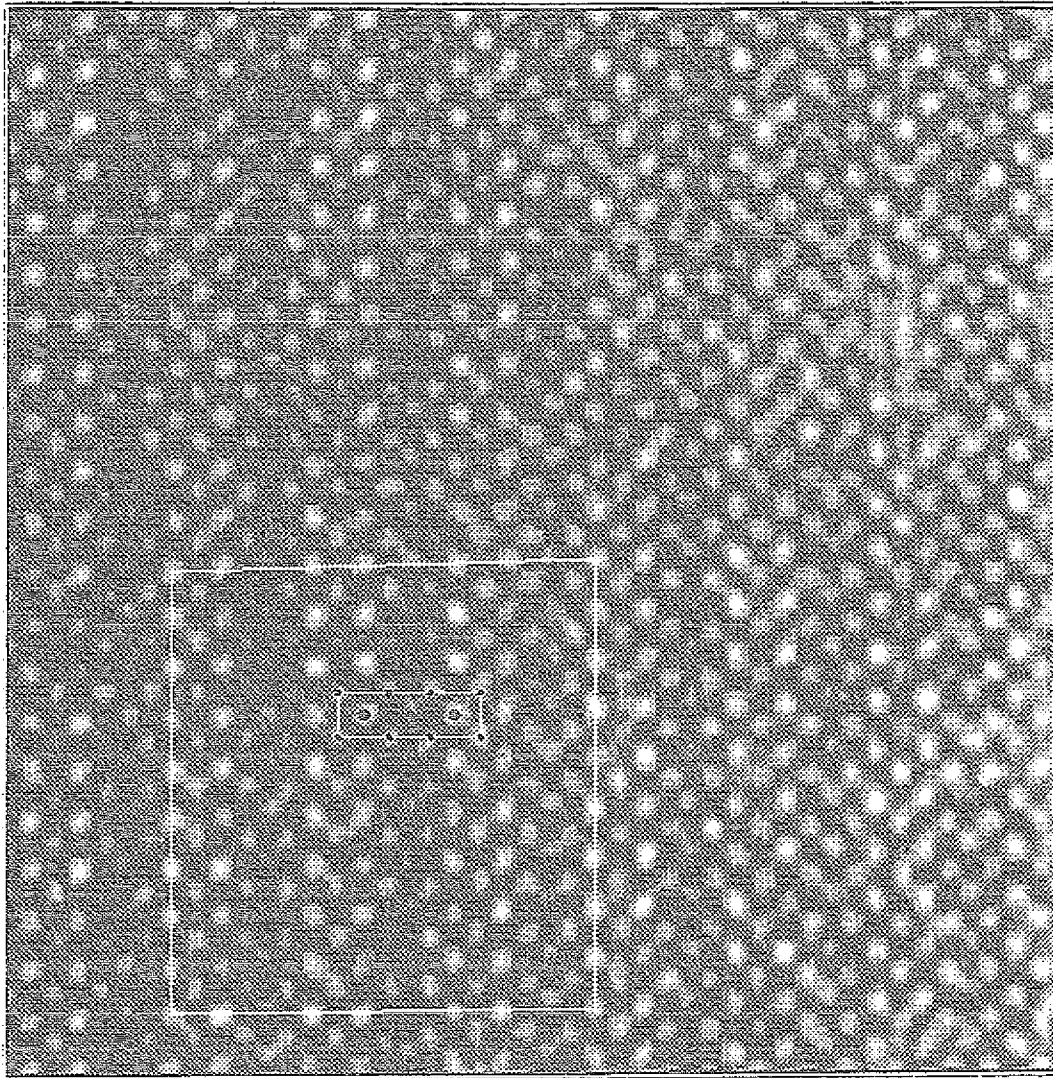
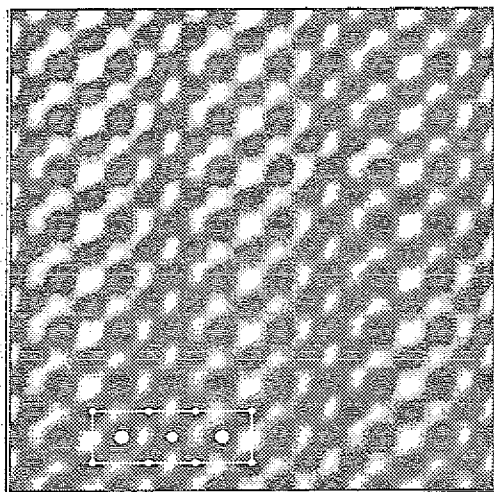
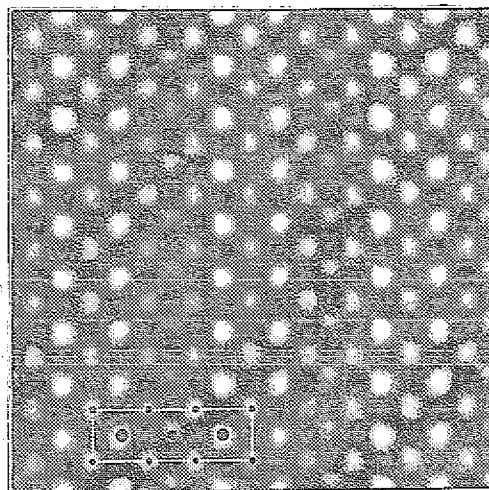


Abbildung 8.10: Hochauflösende Abbildung von $\text{YBa}_2\text{Cu}_3\text{O}_7$ entlang der $[100]$ Zonenachse. Die Abbildung besteht aus 512×512 Bildpunkten. Das zur Rekonstruktion der Wellenfunktion ausgewählte Objektgebiet ist durch einen Rahmen gekennzeichnet. Eine Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ ist innerhalb des Rahmens eingezeichnet. Kleine Kreise bezeichnen Cu-Positionen, der mittlere die Y-Position, und große Kreise die Ba-Positionen der Einheitszelle. Die Kantenlänge des eingezeichneten Rahmens beträgt 3.5 nm.

Experiment

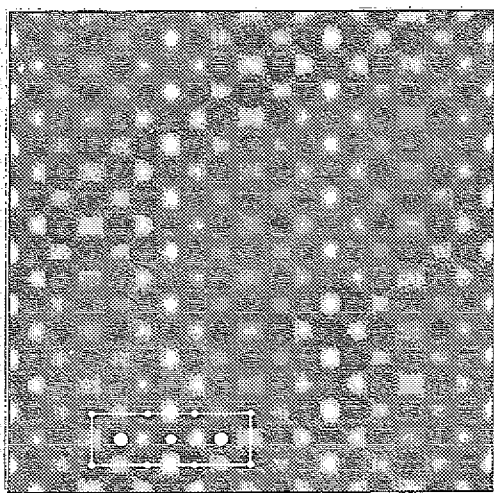


1

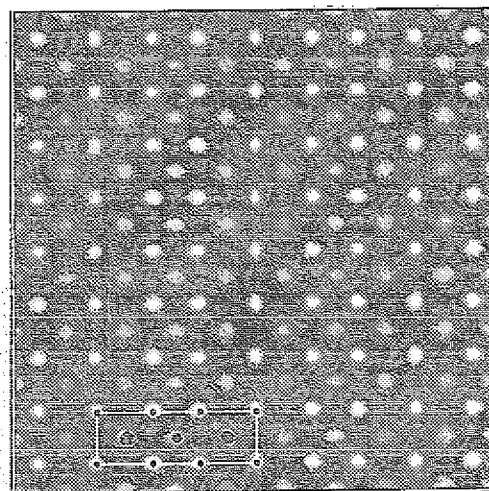


2

Simulation



3



4

Abbildung 8.11: 1: Bildkontrast, der durch Mittelwertbildung des Kontrasts über den in Abb. 8.9 eingerahmten Bereich entsteht. Die projizierten Positionen der Kationen erscheinen dunkel. 2: Bildkontrast, der durch Mittelwertbildung des Kontrasts über den in Abb. 8.10 eingerahmten Bereich entsteht. Die projizierten Positionen der Kationen erscheinen hell. Die Bilder 1 und 2 entsprechen in ihrer Bearbeitungsstufe den in der mittleren Spalte von Abb. 7.6 gezeigten Bildern von Ni_4Mo . 3: Zu Bild 1 am besten korreliertes Bild aus der Bildsimulation. 4: Zu Bild 2 am besten korreliertes Bild aus der Bildsimulation.

dung 8.11 dasjenige simulierte Bild, das bei Variation von Defokus und Objektdicke den höchsten Wert der Kreuzkovarianz zum experimentellen Bild 2 aufweist. Die Simulation von Bild 3 der Abbildung 8.11 basiert auf einer Objektdicke von 3.5 nm und einem Defokuswert von -30 nm, die Simulation von Bild 4 basiert auf derselben Objektdicke und einer Defokussierung von -72 nm.

In Abbildung 8.11 ist zu sehen, daß die simulierten Bilder 3 und 4 bezüglich der Lage der Intensitätsmaxima und Minima recht gut mit den experimentellen Bildern 1 und 2 übereinstimmen. Bei quantitativer Betrachtungsweise der Bildintensität ergeben sich jedoch einige Diskrepanzen. Während im experimentellen Bild 1 globale Intensitätsmaxima an den Tunnelpositionen von Cu1- Ebenen auftreten, liegen im dazu bestkorrelierten simulierten Bild 3 die globalen Intensitätsmaxima an Tunnelpositionen von Yttriumebenen. Ebenso liegen im experimentellen Bild 2 die Stellen höchster Bildintensität an projizierten Ba- Positionen, während im entsprechenden simulierten Bild 4 Cu2- Positionen und Y- Positionen die höchste Intensität aufweisen.

Der Vergleich zwischen experimentellen und simulierten Bildern wurde mit und ohne Symmetriekorrektur der experimentellen Bilder vorgenommen (Siehe Kapitel 7). Da die experimentellen Bilder jedoch nur geringe Abweichungen von der mm Symmetrie der idealen Struktur aufweisen, ergibt die Symmetriekorrektur vor Anwendung des Bildvergleichsverfahrens keine Änderung der ermittelten Defokus- und Dickenwerte. Der Wert der Kreuzkovarianz zwischen der symmetriekorrigierten Version des Bildes 1 und dem simulierten Bild 3 aus Abbildung 8.11 beträgt 0.882, der entsprechende Wert der Kreuzkovarianz zwischen der symmetriekorrigierten Version von Bild 2 und dem simulierten Bild 4 ist 0.884. Diese Werte sind bei weitem geringer als die beim Vergleich zwischen symmetriekorrigierten experimentellen Bildern von Ni_4Mo und entsprechenden simulierten Bildern erzielten Werte um 0.995 (Siehe Kapitel 7).

Betrachtet man diese Diskrepanz vom Standpunkt der Rekonstruktion der Wellenfunktion, so bedeuten Werte der Kreuzkovarianz um 0.88, daß mit den Wellenfunktionen, die mit Hilfe des EMS Pakets [46] berechnet wurden, nur ein minimaler Wert der Evaluationsfunktion (Gl. (8.1)) von etwa 28° erzielt wurde. Es stellt sich die Frage, ob es eine Wellenfunktion gibt, die die experimentellen Bilder 1 und 2 der Abbildung 8.11 besser beschreibt als der aus der Simulation der dynamischen Elektronenbeugung gewonnene Satz von Wellenfunktionen. Eine solche Wellenfunktion sollte einen Wert der Evaluationsfunktion deutlich unter 28° aufweisen.

Für eine Rekonstruktion der Austrittswellenfunktion aus den Bildern 1 und 2 der Abbildung 8.11 müssen die zugehörigen Defokuswerte bekannt sein. Es wird im folgenden davon ausgegangen, daß der durch die Simulation der dynamischen Elektronenbeugung berechnete Satz von Wellenfunktionen zumindest näherungsweise die tatsächlichen Verhältnisse beschreibt und daß die daraus durch Bildvergleich bestimmten Defokuswerte von -30 nm und -72 nm in etwa richtig sind. Für die Rekonstruktion der Austrittswellenfunktion als quantitative Meßmethode ist die Genauigkeit der durch Bildvergleich bestimmten Defokuswerte bei nur mäßiger Übereinstimmung zwischen Bildsimulation und Experiment jedoch unter Umständen nicht ausreichend. Wie an späterer Stelle noch dargelegt wird, kann die hier vorgestellte

Rekonstruktionsmethode herangezogen werden, um den relativen Defokusabstand zweier Bilder vollkommen unabhängig von den Methoden der konventionellen Bildsimulation zu bestimmen.

Um festzustellen, ob mittels des Simulated-Annealing-Algorithmus unter den gegebenen Abbildungsbedingungen die Rekonstruktion der Wellenfunktion aus zwei Bildern möglich ist, wurden die simulierten Bilder 3 und 4 der Abbildung 8.11 als Testbilder herangezogen. Das Ergebnis der auf 53 Strahlen basierenden Rekonstruktion ist in Abbildung 8.12 dargestellt. Auf der linken Seite dieser Abbildung sind Amplitude (oben) und Phase (unten) der rekonstruierten Wellenfunktion dargestellt, auf der rechten Seite ist die Ausgangswellenfunktion, die zur Simulation der Eingabebilder benutzt wurde, gegenübergestellt. Rekonstruierte Wellenfunktion und Ausgangswellenfunktion sind visuell identisch. Der numerische Vergleich von Amplituden und Phasen der beiden Wellenfunktionen, auf dessen Darstellung hier verzichtet wird, ergibt sogar noch geringere Abweichungen als bei der Modellrekonstruktion aus simulierten Abbildungen von Ni_4Mo (Siehe Tabelle 8.1). Dies liegt daran, daß bei einem Objektivblendendurchmesser entsprechend 11 nm^{-1} keiner der an der Abbildung beteiligten Strahlen so stark gedämpft wird, daß sein Beitrag zum Bildkontrast numerisch nicht mehr auflösbar wäre.

Sowohl das Phasen- als auch das Amplitudenmuster der in Abbildung 8.12 dargestellten Wellenfunktion sind unmittelbar mit der entlang $[100]$ projizierten Struktur von $\text{YBa}_2\text{Cu}_3\text{O}_7$ korreliert. Die projizierten Positionen der Yttrium-, Barium- und Kupferatome sind im Amplitudenmuster als lokale Maxima oder Minima zu erkennen, die von ringförmigen Gebieten hoher Amplitude umgeben sind. Im Phasenmuster treten die projizierten Kationenpositionen als stark lokalisierte Maxima über einem gleichförmigen Untergrund niedriger Phase hervor. Die Positionen der schwach streuenden Sauerstoffatome können bei einer Objektdicke von 3.5 nm in der Wellenfunktion nicht beobachtet werden. Obwohl das Objekt bei einer Dicke von 3.5 nm schon nicht mehr als schwaches Phasenobjekt angesehen werden kann, spiegelt das Phasenmuster qualitativ das Kristallpotential wider. Orte hohen Kristallpotentials fallen mit Orten hoher Phase zusammen. Aufgrund der dynamischen Beugungseffekte ist jedoch die Proportionalität von Potential und Phase, die bei einem schwachen Phasenobjekt gilt, nicht mehr gegeben.

Zur Rekonstruktion der Austrittswellenfunktion aus den experimentellen Bildern 1 und 2 der Abbildung 8.11 wurden die durch Vergleich mit simulierten Bildern ermittelten Defokuswerte von -30 nm und -72 nm nur als Anhaltspunkt benutzt. Unter Beibehaltung des Mittelwertes von -51 nm wurden mehrere Rekonstruktionen bei Vorgabe unterschiedlicher relativer Defokusdifferenzen durchgeführt. Um zu überprüfen, ob diese Vorgehensweise prinzipiell geeignet ist, die Defokusdifferenz zweier Bilder zu ermitteln, wurde bei der parallel durchgeführten Testrekonstruktion aus den simulierten Bildern 3 und 4 der Abbildung 8.11 ebenfalls die vorgegebene Defokusdifferenz variiert. Die so erhaltenen minimalen Werte der Evaluationsfunktion sind in Abbildung 8.13 für die Rekonstruktion aus den simulierten Testbildern (oben) und für die Rekonstruktion aus den experimentellen Bildern (unten) dargestellt.

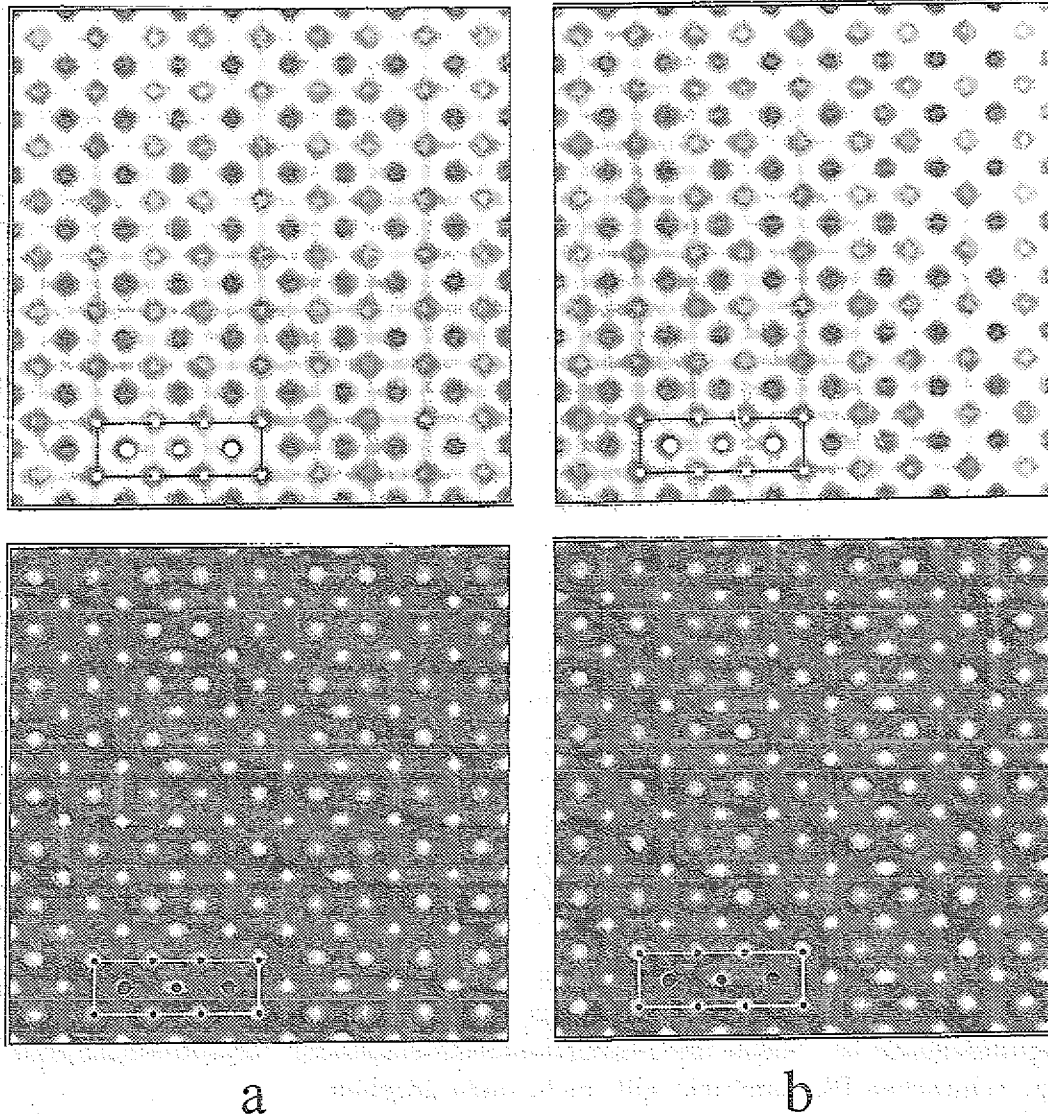


Abbildung 8.12: (a) Aus den simulierten Bildern 3 und 4 der Abbildung 8.11 rekonstruierte Amplitude (oben) und Phase (unten). (b) Amplitude (oben) und Phase (unten) der Wellenfunktion, die als Ausgang zur Simulation der Bilder 3 und 4 in Abbildung 8.11 benutzt wurde. Eine entlang $[100]$ projizierte Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ ist jeweils eingezeichnet. Kleine Kreise markieren Cu-Positionen, mittlere die Y-Position und große Kreise Ba-Positionen. Die Darstellung von Amplitude und Phase basiert auf der Wahl der Nullstrahlphase zu 0° . Die Auflösung der Darstellung ist entsprechend dem zur Bildsimulation benutzten Blendenradius auf 5.5 nm^{-1} begrenzt.

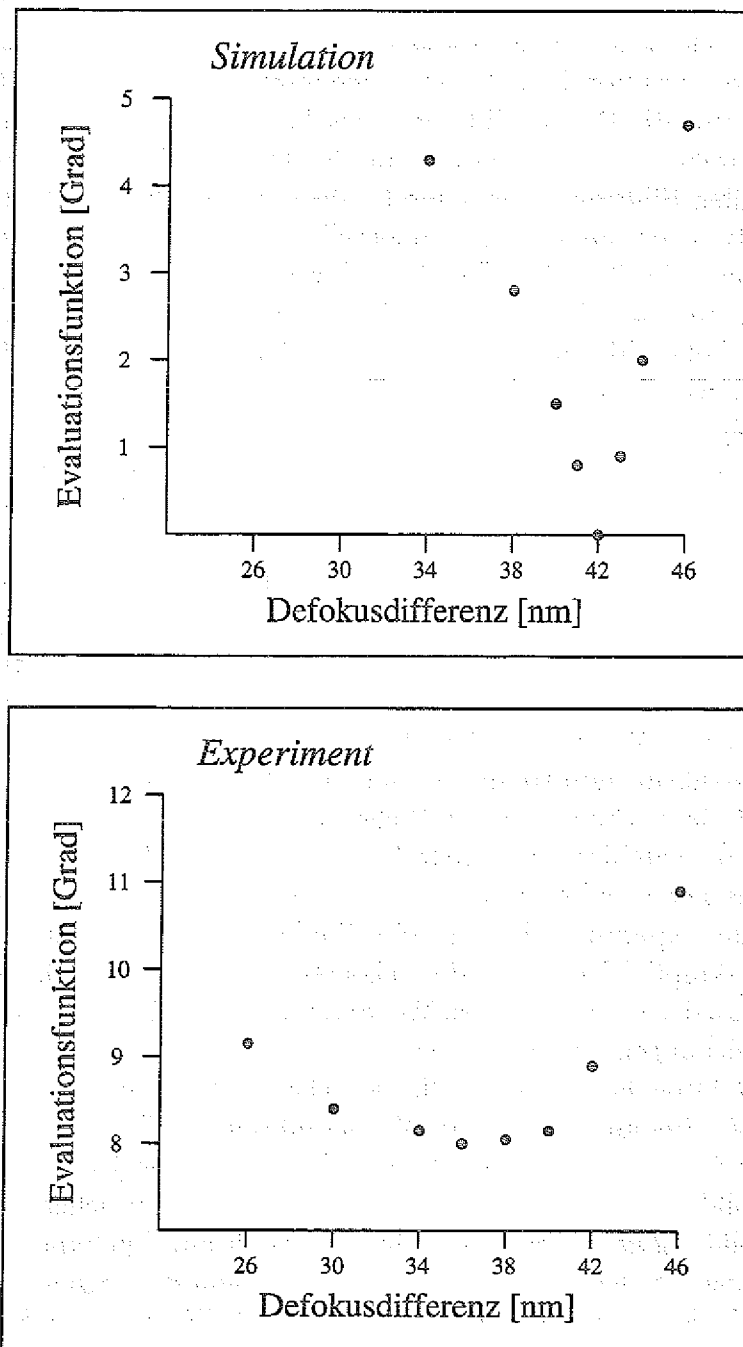


Abbildung 8.13: Auftragung des minimal erzielbaren Wertes der Evaluationsfunktion bei der Rekonstruktion der Wellenfunktion aus zwei Bildern in Abhängigkeit von der vermuteten Defokusdifferenz. Die obere Darstellung bezieht sich auf die Rekonstruktion der Wellenfunktion aus den simulierten Bildern 3 und 4 der Abbildung 8.11. Die untere Darstellung bezieht sich auf die Rekonstruktion der Wellenfunktion aus den experimentellen Bildern 1 und 2 der Abbildung 8.11.

Man sieht in Abbildung 8.13, daß bei der Rekonstruktion aus simulierten Bildern die Vorgabe einer „falschen“ Defokussdifferenz, die nicht genau dem wahren Wert von 42 nm entspricht, sofort einen markanten Anstieg des minimal erzielbaren Wertes der Evaluationsfunktion bewirkt. Dies bedeutet, daß alleine mittels des Rekonstruktionsverfahrens die Defokussdifferenz zwischen den simulierten Eingabebildern eindeutig bestimmbar ist. Zur Bestimmung der Defokussdifferenz ist im Gegensatz zur konventionellen Bildsimulation keinerlei Wissen über die Objektstruktur nötig. Bei der Rekonstruktion aus den experimentellen Bildern ergibt sich ein Minimum der Evaluationsfunktion bei einer Defokussdifferenz von etwa 36 nm. Der bei dieser Defokussdifferenz erzielte minimale Wert der Evaluationsfunktion beträgt 8.0° . Neben dem relativ hohen Minimalwert der Evaluationsfunktion von 8.0° ist auffällig, daß die Evaluationsfunktion zu beiden Seiten des Minimums bei 38 nm bedeutend schwächer ansteigt als im simulierten Modellfall.

Die mit den optimierten Defokusswerten aus den Bildern 1 und 2 der Abbildung 8.11 rekonstruierte Wellenfunktion ist in Abbildung 8.14 dargestellt. Wie im Falle der für eine Objektdicke von 3.5 nm simulierten Wellenfunktion, die in Abbildung 8.12 dargestellt ist, stehen Amplitude und Phase in direktem Zusammenhang mit der entlang [100] projizierten Struktur von $\text{YBa}_2\text{Cu}_3\text{O}_7$. In der Amplitude der Wellenfunktion treten in der Nähe der projizierten Positionen von Yttrium-, Barium- und Kupferatomen deutlich sichtbare lokale Minima auf. Das Phasenmuster weist sehr scharf begrenzte Maxima an den Positionen der Kationen auf und gibt somit die projizierte Struktur unmittelbar wieder. Amplitude und Phase der aus den experimentellen Bildern rekonstruierten Wellenfunktion stimmen mit der simulierten Wellenfunktion in Abbildung 8.12 jedoch nur qualitativ überein. Der markanteste Unterschied zwischen simulierter und rekonstruierter Wellenfunktion besteht in der Tatsache, daß die experimentell ermittelte Wellenfunktion „unschärfer“ ist als die simulierte. Das Amplitudenmuster der rekonstruierten Wellenfunktion weist weniger Details auf als das der simulierten Wellenfunktion und ist zusätzlich von starken Symmetrieabweichungen geprägt. Ebenso sind die Phasenmaxima der rekonstruierten Wellenfunktion bei weitem nicht so scharf an den projizierten Atomorten lokalisiert wie dies bei der simulierten Wellenfunktion der Fall ist.

Wie gut die in Abbildung 8.14 gezeigte Wellenfunktion jedoch die experimentellen Eingabebilder reproduziert, ist in Abbildung 8.15 zu sehen. Die im oberen Teil dieser Abbildung gezeigten Bilder sind identisch mit den experimentellen Eingabebildern 1 und 2 aus Abbildung 8.11. Darunter sind diejenigen Bilder zu sehen, die auf der Basis der in Abbildung 8.14 gezeigten Wellenfunktion für die Defokusswerte -32 nm und -70 nm berechnet wurden. Die experimentellen Eingabebilder und die aus der rekonstruierten Wellenfunktion berechneten Bilder unterscheiden sich optisch nur in geringfügigen Kontrastdetails. Der Wert der Evaluationsfunktion von 8° , der diese Unterschiede beschreibt, ist bei weitem geringer als der Bestwert von 28° , der beim Vergleich der experimentellen Eingabebilder mit den simulierten Bildern 1 und 2 der Abbildung 8.11 gefunden wurde.

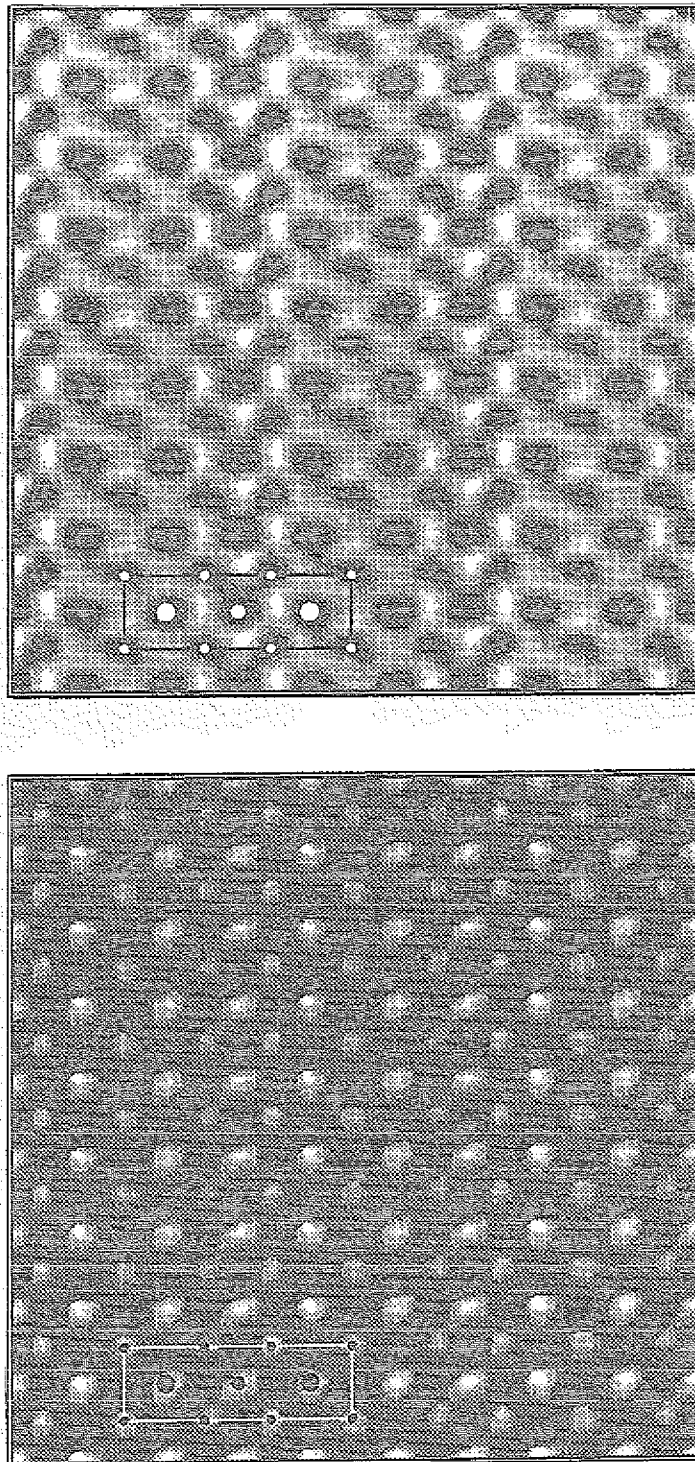


Abbildung 8.14: Aus den experimentellen Bildern 1 und 2 der Abbildung 8.11 rekonstruierte Amplitude (oben) und Phase (unten). Die Darstellung von Amplitude und Phase basiert auf der Wahl der Nullstrahlphase zu 0° . Die Auflösung der Darstellung ist entsprechend dem experimentellen Blendenradius auf 5.5 nm^{-1} begrenzt. Eine entlang $[100]$ projizierte Einheitszelle von $\text{YBa}_2\text{Cu}_3\text{O}_7$ ist jeweils eingezeichnet.

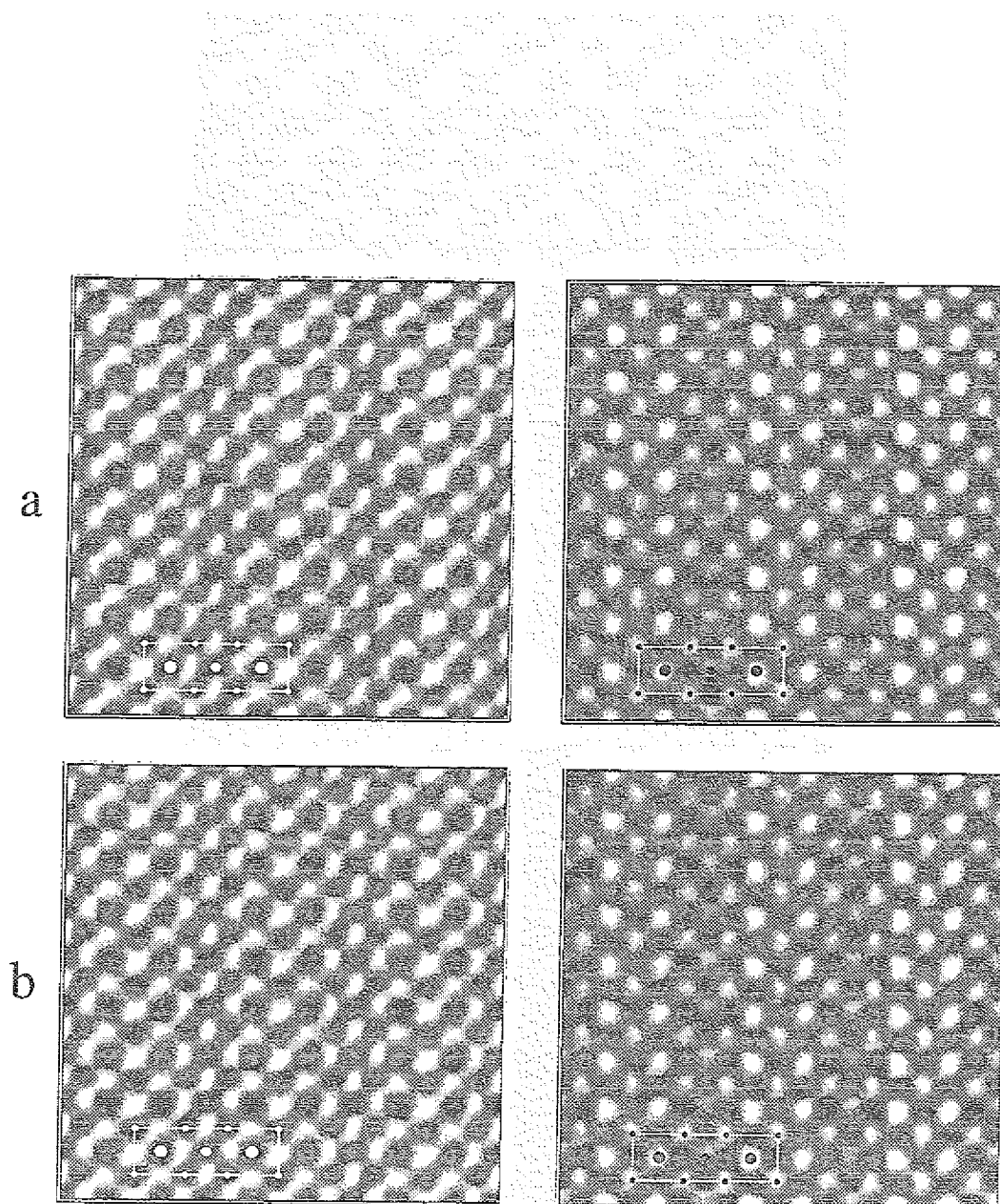


Abbildung 8.15: Vergleich zwischen experimentellen Ausgangsbildern zur Rekonstruktion der Wellenfunktion (a) und den aus der rekonstruierten Wellenfunktion wiederum berechneten Bildern (b). Die Ausgangsbilder sind mit den in Abbildung 8.11 gezeigten Bildern 1 und 2 identisch.

8.8 Diskussion

Durch die Möglichkeit, den Aufwand für die Bildberechnung mittels der „*differential updating*“ Technik für den Simulated-Annealing-Algorithmus maßgeschneidert zu optimieren, ergibt sich bei Rekonstruktion von etwa 50 Fourierkomponenten gegenüber dem hier benutzten genetischen Algorithmus ein Rechengeschwindigkeitsvorteil von etwa einer Größenordnung. Bei Rekonstruktionen mit weit höheren Strahlzahlen würde der Vorteil noch dramatischer ausfallen, da beim *simulated annealing* ein Aufwand proportional der Strahlzahl N gegenüber einem Aufwand proportional N^2 beim genetischen Algorithmus steht. Zusätzlich sind die Parameter, die beim *simulated annealing* den Temperaturfahrplan festlegen, völlig problemlos handzuhaben, während die Steuerung des Suchverhaltens beim genetischen Algorithmus über Populationsgröße und Mutationsrate vergleichsweise schwieriger ist. Somit ist die in dieser Arbeit zum Zwecke der Rekonstruktion der Austrittswellenfunktion benutzte Implementation des Simulated-Annealing-Algorithmus eindeutig der benutzten Implementation eines genetischen Algorithmus vorzuziehen. Die in diesem Abschnitt gezeigten Rekonstruktionen wurden deshalb mit dem Simulated-Annealing-Algorithmus durchgeführt. Parallel durchgeführte Vergleiche der Genauigkeit der erzielten Resultate sowie der Anzahl der benötigten Evaluationen zeigen jedoch, daß die Sucheeffizienz des genetischen Algorithmus dem Simulated-Annealing-Algorithmus ebenbürtig ist. Der Vorteil des Simulated-Annealing-Algorithmus beruht also hauptsächlich auf der problemspezifischen Anpaßbarkeit der benutzten Evaluationsfunktion und nicht auf einer besseren Sucheeffizienz des Algorithmus selbst.

Die in diesem Abschnitt gezeigte Rekonstruktion, bei der die simulierten Testbilder 3 und 4 aus Abbildung 8.11 als Eingabe dienten, demonstriert eindrucksvoll, daß mittels Simulated-Annealing-Algorithmus auch die Wellenfunktion sehr komplex aufgebauter Einheitszellen „perfekt“ rekonstruiert werden kann. Der Algorithmus nutzt bis hinunter zur numerischen Auflösungsgrenze der verwendeten Zahlendarstellung jedes Quantum an Bildinformation und liefert Ergebnisse, deren Genauigkeit nur noch von analytisch errechneten Resultaten übertroffen werden könnte. Die zur Rekonstruktion periodischer Wellenfunktionen benötigte Rechenzeit gestattet den Einsatz in interaktiven Computersitzungen und bedarf kaum noch einer Beschleunigung. Zusätzlich ist der Algorithmus sehr „sparsam“ bezüglich der benötigten Bildinformation. Es werden nur einige wenige Eingabebilder gebraucht, was den experimentellen Aufwand für eine Rekonstruktion sehr vereinfacht. Ein weiterer Vorteil der Methode besteht darin, daß keinerlei Annahmen über die Stärke der Streuung gemacht werden müssen und somit keinerlei Anlehnungen an die Näherungen des schwachen Phasenobjekts und des linearen Kontrasttransfers nötig sind. Die Problemstellung, die Austrittswellenfunktion periodischer Objekte unter vollständiger Einbeziehung des nichtlinearen Kontrasttransfers sowie unter Berücksichtigung der partiellen Kohärenz der Elektronenquelle zu rekonstruieren, kann somit als zufriedenstellend gelöst betrachtet werden.

Die Rekonstruktion aus den experimentellen Bildern 1 und 2 der Abbildung 8.11 zeigt jedoch einige Probleme auf, die nicht auf der methodisch-analytischen Seite

sondern auf der Seite der experimentellen Datenerfassung zu suchen sind. Wie im vorangehenden Absatz bereits geschildert wurde, können mit den hier verwendeten nicht-deterministischen Rekonstruktionsalgorithmen beinahe beliebig genaue Resultate erzielt werden, sofern die Eingabedaten konsistent sind. Diese Eingabedaten lassen sich in drei Gruppen aufteilen: Daten, die das Objekt beschreiben, Parameter die das Mikroskop beschreiben sowie die Eingabebilder selbst.

Bezüglich des Objekts wird als alleiniges Wissen der Satz der am Beugungsbild beteiligten reziproken Gittervektoren vorausgesetzt. Diese Gittervektoren können problemlos und mit ausreichender Genauigkeit aus einem Beugungsbild bestimmt werden. Bezüglich des Mikroskops gehen die Wellenlänge des Elektronenstrahls λ , die sphärische Aberrationskonstante C_s , der Halbwinkel der Strahlkonvergenz θ_c , die Defokusstreuung Δ sowie die zu den Eingabebildern gehörenden Defokuswerte Z_i ein. Die Wellenlänge λ ist mit ausreichender Genauigkeit bekannt, die Parameter θ_c und Δ sind ebenfalls mit einer Genauigkeit von etwa $\pm 20\%$ bekannt. Diese Unsicherheit beeinflusst den Bildkontrast nur in sehr geringem Ausmaß. Die sphärische Aberrationskonstante C_s sollte bei den hier durchgeführten Rekonstruktionen bis zu einer Auflösung von 0.17 nm mit einer Unsicherheit, die kleiner als 5 % ist, bekannt sein. Die zur Rekonstruktion benutzte Herstellerangabe für C_s erfüllt dieses Kriterium mit hoher Wahrscheinlichkeit. Die relativen Defokuswerte lassen sich mit dem Algorithmus selbst optimieren, der absolute Defokuswert hat auf das Konvergenzverhalten keinen Einfluß. Hat man einen falschen absoluten Defokuswert Z_0 als Eingabe für die Rekonstruktion gewählt, kann die so ermittelte „falsche“ Wellenfunktion $\Psi(\vec{k}, Z_0)$ durch einfache Propagation um die absolute Defokusdifferenz ΔZ korrigiert werden, falls die absolute Defokusdifferenz bekannt ist. Die „richtige“ Wellenfunktion $\Psi(\vec{k}, Z_0 + \Delta Z)$ lautet dann

$$\Psi(\vec{k}, Z_0 + \Delta Z) = \Psi(\vec{k}, Z_0) \cdot \exp \left[-i\pi\lambda\Delta Z\vec{k}^2 \right]. \quad (8.7)$$

In der obigen Gleichung wurde der Einfluß des Dämpfungsterms E_s vernachlässigt (Siehe Gl. (5.70) sowie Anhang B). Dieser Term ändert sich jedoch für Defokusdifferenzen von $\Delta Z = \pm 10$ nm kaum merklich. Bei einer signifikanten Änderung des Dämpfungsterms E_s mit der Defokussierung wäre sonst auch der absolute Defokuswert aus Hochauflösungsaufnahmen bestimmbar. Dies bedeutet, daß eine falsche Wahl des Defokuswertes die Konsistenz der Eingabedaten kaum beeinträchtigt.

Die dem Algorithmus übergebenen experimentellen Daten müssen jedoch partiell inkonsistent sein, da der minimale Wert der Evaluationsfunktion 8° beträgt anstatt 0° , wie dies im Falle vollständig konsistenter Daten bei der Rekonstruktion simulierter Bilder der Fall ist. Nachdem nun die beiden Sätze von Eingabedaten, nämlich die reziproken Gittervektoren des Objekts sowie die optischen Parameter, als Inkonsistenzquellen mit hoher Wahrscheinlichkeit ausgeschlossen werden können, kommen nur die Eingabebilder selbst als Quelle für widersprüchliche Information in Frage.

Die Fehlerquellen bei der experimentellen Datenaquisition sind leicht einsichtig. Die Aufnahme der Bildintensität auf photographischen Filmen mit den bereits diskutierten Unabwägbarkeiten der Korrektur der Filmempfindlichkeit sowie des Modulationstransfers des Films können die Konsistenz der Eingabedaten beeinträch-

tigen. Ebenso führt eine mögliche gegenseitige Verschiebung der auf den beiden Aufnahmen zur Rekonstruktion ausgewählten Objektgebiete zu Diskrepanzen, falls die Objektdicke nicht konstant ist.

Eine weitere Fehlerquelle besteht in der Möglichkeit daß die Objektorientierung in den beiden Aufnahmen nicht genau dieselbe ist. In beiden Eingabebildern der Abbildung 8.11 kann man eine leichte Fehlorientierung des Objekts erkennen. In Bild 1 ist zu erkennen, daß die Kontrastmaxima entlang der $[013]$ Kristallrichtung, die in der Abbildung parallel zur Bilddiagonalen von links unten nach rechts oben verläuft, nicht so scharf aufgelöst sind, wie entlang der symmetrieäquivalenten senkrecht dazu verlaufenden Richtung. Dieses Verhalten spiegelt sich in der Phase der rekonstruierten Wellenfunktion, die in Abbildung 8.14 dargestellt ist, wider. Die Phasenmaxima sind entlang der im Bild von links unten nach rechts oben verlaufenden $[013]$ Kristallrichtung etwas „verschmiert“. Analog kann man in Bild 2 der Abbildung 8.11 erkennen, daß das Intensitätsmaximum, das sich idealerweise an der projizierten Yttriumposition befinden sollte, nicht zentrisch zwischen den vier benachbarten Cu²⁺-Positionen liegt. In der Phase der rekonstruierten Wellenfunktion weicht wahrscheinlich deshalb das Maximum an der Yttriumposition von der idealerweise kreisförmigen Form ab und ist ellipsenähnlich gestreckt.

Da die in den Bildern 1 und 2 der Abbildung 8.11 beobachteten leichten Asymmetrien wahrscheinlich nicht auf dieselbe Fehlorientierung des Objekts zurückzuführen sind, enthalten diese Bilder inkonsistente Information. Es ist in diesem Fall auch nicht möglich, alle Fourierwellen der Bildintensität in den Bildern 1 und 2 phasengleich zur Deckung zu bringen. Das Umschlagen der Phase von 180° (weiß) auf -180° (schwarz) in Abbildung 8.14, das an der in Bezug auf die markierte Einheitszelle rechts liegenden Bariumposition beobachtet werden kann, ist auf diesen Effekt zurückzuführen. Verschiebt man vor der Rekonstruktion eines der beiden Bilder minimal gegenüber dem anderen, tritt ein ähnlicher Effekt an einer anderen Stelle der rekonstruierten Phase auf.

Die bisher genannten Quellen für widersprüchliche Bildinformation können bei digitaler Aufnahme der Bilder mittels einer CCD- (*charge coupled device*) Kamera weitgehend ausgeschaltet werden [68, 6, 69]. Neben dem Vorteil der linearen Aufzeichnung der Elektronenintensität innerhalb eines großen dynamischen Bereichs werden durch die Nutzung einer CCD- Kamera mechanische Instabilitäten beim Filmtransport vermieden, die die Probenorientierung zwischen zwei Aufnahmen unkontrollierbar beeinflussen können. Ein weiterer Vorteil der Nutzung einer CCD- Kamera besteht darin, daß das Auffinden exakt ein und derselben Probenstelle auf verschiedenen Aufnahmen, das auf Filmplatten kaum möglich ist, durch die Anwendung von Kreuzkorrelationstechniken fast automatisch erledigt werden kann. Ebenso können bei Benutzung einer CCD- Kamera bei weitem kürzere Belichtungszeiten gewählt werden als dies bei photographischen Filmen möglich ist. Auf diese Weise wird der Auflösungsverlust, der durch eine mögliche Drift des Objekts verursacht wird, stark reduziert.

Nachdem nun mögliche Fehlerquellen bei der Bildaufnahme diskutiert wurden, stellt sich die Frage nach der Relevanz der rekonstruierten Wellenfunktion, die in Abbildung 8.14 dargestellt ist. Akzeptiert man die in der Abbildung 8.11 dargestell-

ten Bilder 1 und 2, die aus der Mittelwertbildung über 27 Einheitszellen gewonnen wurden, als repräsentativ für die in den Abbildungen 8.9 und 8.10 dargestellten „Rohbilder“, so ist die experimentell rekonstruierte Wellenfunktion von weit höherer Relevanz als die durch Simulation der dynamischen Elektronenbeugung berechnete Wellenfunktion, die in Abbildung 8.12 dargestellt ist. Dies läßt sich durch einfachen optischen Vergleich der experimentellen Eingabebilder und den aus den beiden Wellenfunktionen berechneten Bildern feststellen. Die in der Abbildung 8.15 gezeigte Gegenüberstellung zwischen den experimentellen Eingabebildern und den aus der rekonstruierten Wellenfunktion berechneten Bildern ist bei weitem überzeugender als die in Abbildung 8.11 gezeigte Gegenüberstellung zwischen den experimentellen Eingabebildern und den aus der simulierten Wellenfunktion berechneten Bildern. In Abbildung 8.15 können nur äußerst geringfügige Abweichungen zwischen den Eingabebildern und den aus der rekonstruierten Wellenfunktion berechneten Bildern festgestellt werden.

Der optische Eindruck läßt sich quantitativ durch die entsprechenden Werte der Evaluationsfunktion untermauern. Der Wert von 28° , der durch die simulierte Wellenfunktion erreicht wird, deutet auf eine geringe Relevanz dieser Wellenfunktion hin. Hingegen ist der Wert von 8° , der die rekonstruierte Wellenfunktion charakterisiert, von ähnlicher Relevanz wie der Wert von 3.8° , der bei der Rekonstruktion im Ni_4Mo System erzielt wurde. In Anbetracht der beinahe doppelten Strahlzahl von 53 gegenüber 29 Strahlen bei Ni_4Mo und in Anbetracht der Tatsache, daß auf jegliche Symmetriekorrektur verzichtet wurde, ist der erzielte Wert von 8° sehr plausibel.

Ein charakteristischer Unterschied zwischen den Ergebnissen der Bildsimulation und den experimentellen Aufnahmen besteht in der Auflösung der Bilder. Dieser Unterschied läßt sich schon durch einfachen optischen Vergleich der in Abbildung 8.11 gegenübergestellten Bilder erkennen. So sind zum Beispiel die Intensitätsmaxima an den Cu- und Y- Positionen des simulierten Bildes 4 bei weitem schärfer abgebildet als die entsprechenden Maxima im experimentellen Bild 2.

Dieser Unterschied zwischen Experiment und Simulation läßt sich deutlich anhand der Diagramme der Abbildung 8.13 nachweisen, in denen der minimal erzielbare Wert der Evaluationsfunktion über der angenommenen Defokusdifferenz aufgetragen ist. Bei den simulierten Bildern „verrät“ sich die wahre Defokusdifferenz von 42 nm durch ein sehr markant ausgeprägtes Minimum an dieser Stelle. Die Existenz eines so stark lokalisierten Minimums ist nur möglich, wenn in beiden Eingabebildern Strahlen hoher Raumwellenlänge konsistent zum Bildkontrast beitragen. Im Falle hoher Auflösung ändert sich der Bildkontrast bei geringer Defokusänderung bedeutend stärker als wenn vorwiegend Strahlen mit niedriger Auflösung eine Rolle spielen, da wegen $\partial\chi/\partial Z \sim k^2$ eine quadratische Abhängigkeit zwischen defokusinduzierter Kontraständerung und Raumwellenlänge besteht (Siehe Gln. 5.69 und 5.70). Das sehr flach verlaufende Minimum, das im Falle der experimentellen Bilder entsteht, ist neben der schon diskutierten partiellen Inkonsistenz der Bildinformation auch auf das Fehlen höherer Raumfrequenzen zurückzuführen.

Die Ursachen für die im Vergleich zur Simulation unerwartet schlechte experimentelle Auflösung können sowohl auf der Seite des Objekts als auch auf der Seite der elektronenoptischen Abbildung zu suchen sein. Bei der Simulation der Abbildung

wird normalerweise der Einfluß mechanischer Vibrationen während der Aufnahme nicht berücksichtigt. Solche Vibrationen beeinträchtigen die erzielte Auflösung empfindlich. Es ist jedoch nicht möglich, das Vorhandensein solcher Vibrationen im nachhinein anhand der Bilddaten nachzuweisen und von Effekten zu separieren, die sich schon *vor* der Abbildung in der Wellenfunktion manifestieren. Solche Effekte, die schon bei der dynamischen Elektronenbeugung eine Rolle spielen können, werden im folgenden kurz diskutiert.

Bei der Simulation der Wellenfunktion – sei es mittels Blochwellenverfahren oder sei es mittels Multislice-Verfahren – werden üblicherweise Annahmen über das Objekt gemacht, die nicht in jedem Falle zutreffend sind. So wird angenommen, daß die einfallende Elektronenwelle eine nahezu ebene Welle ist, die nur durch eine Strahlkonvergenz von etwa einem Millirad aufgefächert ist. Befindet sich amorphes Material an der Probenoberfläche, so ist der Raum über der Eintrittsfläche jedoch nicht mehr potentialfrei und die einfallende Welle weicht von der angenommenen Form bereits ab. Ebenso kann sich an der Unterseite der Probe eine Schicht aus amorphem Material befinden. Die Annahme der Propagation der Welle von Objektebene zu Bildebene durch potentialfreien Raum, die durch den Propagationsterm der Aberrationsfunktion (Gl. (5.8)) beschrieben wird, ist dann ebenfalls nicht zutreffend. Die elektrostatischen Potentiale der in den amorphen Deckschichten befindlichen Atome wirken wie verzerrende Linsen auf das Wellenfeld. Untersuchungen zeigen, daß die letztendlich bildverzerrende Wirkung amorpher Deckschichten erhebliches Ausmaß haben kann [82]. Die Tatsache, daß sich benachbarte Einheitszellen in den Aufnahmen der Abbildungen 8.9 und 8.10 nie genau gleichen, läßt sich so erklären.

Amorphe Schichten an Probenober- und unterseite bilden sich durch Schädigung des Materials durch den Beschuß mit Ar^+ Ionen bei der Probenpräparation oder durch Schädigung des Materials durch den 400 kV- Elektronenstrahl unmittelbar bei der Beobachtung. Viele Versuche des Verfassers, eine Defokus-Serie von $\text{YBa}_2\text{Cu}_3\text{O}_7$ anzufertigen, mißlangen, weil das Material durch die Einwirkung des Elektronenstrahls völlig zerstört wurde und nach wenigen Minuten keinerlei Periodizität mehr aufwies. So ist zu vermuten, daß nicht nur amorphe Deckschichten den Bildkontrast verschlechtern, sondern daß auch das verbleibende Material nicht mehr dem Strukturmodell entspricht, das als Ausgang für die Simulation der dynamischen Elektronenbeugung herangezogen wird.

Bei der Simulation der dynamischen Elektronenbeugung wird üblicherweise eine perfekte Periodizität der Struktur vorausgesetzt. Rechnungen mit der Blochwellenmethode für entlang der [100] Zonenachse orientiertes $\text{YBa}_2\text{Cu}_3\text{O}_7$ zeigen, daß die Lösung für die Wellenfunktion erst ab einer Strahlzahl von etwa 400 konvergiert. Bei dieser Strahlzahl werden im Kristallpotential Fourierkoeffizienten mit Raumwellenlängen bis 0.05 nm berücksichtigt. Es ist wahrscheinlich, daß wegen der bei zunehmender Strahlenschädigung stattfindenden Atomverlagerungen auf Zwischengitterplätze zuerst die Fourierkoeffizienten des Kristallpotentials empfindlich gedämpft oder eliminiert werden, die hohe Auflösung besitzen und die bei der Rechnung am idealen periodischen Gitter eine wichtige Rolle spielen.

Neben diesen eher statistisch verteilten Auslenkungen der Atompositionen um ihre Ideallage sind auch systematische Auslenkungen in den Abbildungen 8.9

und 8.10 zu erkennen. An der Tatsache, daß jeweils gegenüberliegende Kanten der in diesen Abbildungen eingezeichneten Rahmen keineswegs parallel verlaufen, erkennt man, daß die abgebildete $\text{YBa}_2\text{Cu}_3\text{O}_7$ -Schicht unter starken Spannungen steht. Da diese Spannungen schon mit „bloßem Auge“ erkennbar sind, haben die Abweichungen der Atompositionen von ihren Ideallagen erhebliches Ausmaß und können wahrscheinlich ebensowenig bei der Berechnung der Wellenfunktion vernachlässigt werden. Nimmt man an, daß die Verspannungen in einer Ebene senkrecht zur Bildebene ebenso stark sind, daß sie mit „bloßem Auge“ erkennbar wären, so ist die in der Simulation angenommene Säulenordnung der Atome parallel zur Strahlrichtung empfindlich gestört. Allein bei geometrischer Projektion der Struktur entlang der Strahlrichtung tritt in solch einem Falle schon ein erheblicher Auflösungsverlust ein.

8.9 Ausblick

Die im Rahmen dieser Arbeit entwickelte Methode zur Rekonstruktion der Austrittswellenfunktion mittels nicht-deterministischer Algorithmen besitzt Leistungsmerkmale, die in mancher Hinsicht als nahezu perfekt eingestuft werden können. Hierzu zählt die Allgemeingültigkeit der Methode, da keinerlei Restriktionen bezüglich des Grades der Nichtlinearität der Abbildung nötig sind. Des weiteren ist die Genauigkeit der erzielten Resultate nicht von der Methode selbst, sondern ausschließlich von der Qualität der Eingabedaten abhängig. Bei Konsistenz der Eingabedaten können Resultate nahezu beliebiger Genauigkeit erzielt werden. Ein dritter, nicht zu unterschätzender Vorteil besteht in der Tatsache, daß nur wenige Eingabebilder zur Rekonstruktion benötigt werden. Dadurch kann der experimentelle Aufwand in Grenzen gehalten werden.

Der – soweit es der Autor übersehen kann – einzige Nachteil der Methode besteht in den langen Rechenzeiten, die für eine Rekonstruktion in Kauf genommen werden müssen. Als die ersten Versuchsprogramme konzipiert wurden, um die Tauglichkeit der Methode an sich zu überprüfen, kristallisierte sich dieser Nachteil schnell heraus. Die ersten Rekonstruktionen wurden auf einer „workstation“ vorgenommen, deren Rechengeschwindigkeit aufgrund der rasanten technischen Entwicklung auf dem Hardwaresektor der Rechengeschwindigkeit eines Heimcomputers der jüngsten Generation bereits unterlegen ist. Rechenzeiten von einem ganzen Tag waren keine Seltenheit. Die in dieser Arbeit vorgenommene Beschränkung der Rekonstruktionen auf periodische Bildmotive ist eine Folge dieses ersten Eindrucks.

Mit der Möglichkeit, die Rekonstruktionsprogramme zeitweise auf einer „workstation“ der neuesten Generation erproben zu können, zeichnete sich eine wesentliche Entschärfung des Rechenzeitproblems ab. Durch den Umstieg auf die neue Rechenplattform konnte eine Geschwindigkeitsgewinn von etwa einem Faktor 40 erzielt werden. Die parallel dazu gemachten Anstrengungen, die anfängliche Version des Simulated-Annealing-Algorithmus programmtechnisch zu optimieren, erbrachte nochmals – je nach zu lösender Strahlzahl – eine Effizienzsteigerung von etwa einer Größenordnung.

Nimmt man die Anzahl der Fourierkomponenten als Maß, so ist die Leistungsfähigkeit der gegenwärtigen „periodischen“ Implementation des Rekonstruktionsverfahrens mit der des Blochwellenverfahrens zur Berechnung der Wellenfunktion vergleichbar. Aus Gründen der Recheneffizienz und der numerischen Stabilität wird die Nutzung des Blochwellenverfahrens meist auf die Behandlung von einigen hundert Strahlen beschränkt. Da jedoch im allgemeinen nur ein Bruchteil der zur Berechnung der Austrittswellenfunktion nötigen Strahlen tatsächlich an der Abbildung teilnimmt, sollte auch jede mit dem Blochwellenverfahren berechenbare Wellenfunktion mittels des hier vorgestellten Verfahrens problemlos im Rahmen der instrumentellen Auflösung rekonstruierbar sein.

So ist es jetzt schon möglich, die Wellenfunktion komplex aufgebauter periodischer Strukturen, wie am Beispiel von $\text{YBa}_2\text{Cu}_3\text{O}_7$ demonstriert wurde, im Zeitrahmen interaktiver Computersitzungen zu ermitteln. Mit Rechenzeiten von wenigen Minuten, wie sie für die hier vorgeführte Lösung eines 53-Strahl Problems benötigt werden, ist die Methode jedoch bei weitem noch nicht an der Grenze ihrer Leistungsfähigkeit angelangt. Wahrscheinlich können Probleme, bei denen weit über hundert oder sogar einige hundert Fourierkomponenten eine Rolle spielen, auf einer schnellen Rechenanlage in annehmbarer Zeit gelöst werden. Hier bleibt die interessante Frage zu klären, welche Strahlzahlen noch in annehmbaren Rechenzeiten zu bewältigen sind und ob somit auch nicht-periodische Objekte mit der vorgestellten Methode rekonstruierbar sind.

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit quantitativen Methoden zur Auswertung des Bildkontrastes hochaufgelöster elektronenmikroskopischer Aufnahmen sowie mit Methoden zur Rekonstruktion der quantenmechanischen Wellenfunktion an der Austrittsfläche von kristallinen Beobachtungsobjekten.

Es wurde ein schnelles numerisches Bildvergleichsverfahren entwickelt, das es ermöglicht, Aufnahmen periodischer Objekte mit solchen Bildern zu vergleichen, die mittels Simulation von dynamischer Elektronenbeugung und elektronenoptischer Abbildung berechnet wurden. Das Verfahren gestattet es, innerhalb kürzester Zeit typischerweise 10 000 Abbildungen zu simulieren, die den gesamten Wertebereich der im Experiment meist unbekannten Parameter *Objektdicke* und *Defokus* überdecken. Dabei werden die simulierten Bilder mit experimentellen Eingabebildern verglichen und die Übereinstimmung in Form des Kreuzkovarianzkoeffizienten dargestellt. Anhand einer experimentellen Defokus-Serie der intermetallischen Legierung Ni_4Mo wird demonstriert, daß mit diesem Verfahren eine eindeutige und präzise Bestimmung der experimentellen Abbildungsparameter Objektdicke und Defokus möglich ist.

Die Eignung nicht-deterministischer Algorithmen zur Rekonstruktion der quantenmechanischen Wellenfunktion an der Austrittsfläche kristalliner Präparate wurde untersucht. Dabei kommt eine modifizierte Version der oben beschriebenen schnellen Bildvergleichstechnik zum Einsatz. Die wichtigsten Ergebnisse dieser Untersuchungen sind:

- Sowohl Algorithmen, die auf dem Prinzip des *simulated annealing* basieren, als auch genetische Algorithmen sind zur Rekonstruktion der Austrittswellenfunktion periodisch aufgebauter Objekte bestens geeignet.
- Nichtlineare Abbildungsbedingungen sowie die Auswirkungen von nur partiell kohärenter Einstrahlung können in vollem Maße berücksichtigt werden. Es müssen keinerlei Annahmen oder Einschränkungen bezüglich des Grades der Nichtlinearität der Abbildung gemacht werden.
- Für eine Rekonstruktion sind als Eingabe meist nur zwei, unter unterschiedlichen Defokussierungsbedingungen aufgenommene Aufnahmen erforderlich. In Hinblick auf eine Verbesserung des Verhältnisses von Signal zu Rauschen kann jedoch die Verwendung zusätzlicher Bilder von Vorteil sein.
- Mit dem hier vorgestellten Rekonstruktionsverfahren gelang die Rekonstruktion einer periodischen Austrittswellenfunktion unter idealisierten Bedingungen sogar aus einem einzigen Eingabebild.

- Die Genauigkeit der bei der Rekonstruktion erzielten Ergebnisse ist von beinahe analytischer Qualität und hängt nur von der Qualität der Eingabebilder und der Kenntnis der instrumentellen Parameter ab.
- Mittels der genannten Algorithmen können die relativen Defokusabstände experimenteller Aufnahmen ohne Zuhilfenahme eines Strukturmodells bestimmt werden. Ebenso kann die Konsistenz der in zwei Abbildungen enthaltenen Bildinformation überprüft und quantifiziert werden.
- Die hier eingesetzte Version eines Simulated-Annealing-Algorithmus ist aufgrund der besseren problemspezifischen Anpassungsmöglichkeit in seiner numerischen Effizienz der untersuchten Version eines genetischen Algorithmus eindeutig überlegen.
- Mittels des Simulated-Annealing-Algorithmus ist die Rekonstruktion einer aus 50 Strahlen bestehenden Austrittswellenfunktion innerhalb weniger Minuten durchführbar. Stochastische Algorithmen scheinen aus Rechenzeitgründen jedoch weniger gut für die Rekonstruktion nichtperiodischer Bildmotive geeignet zu sein, die die Einbeziehung ungleich höherer Strahlzahlen erforderlich machen.
- Die Rekonstruktion der Austrittswellenfunktion aus experimentellen Bildern der intermetallischen Legierung Ni_4Mo ergibt eine sehr gute Übereinstimmung mit Simulationsergebnissen. Im Gegensatz dazu wurden bei der Rekonstruktion der Austrittswellenfunktion aus experimentellen Bildern des Hochtemperatursupraleiters $\text{YBa}_2\text{Cu}_3\text{O}_7$ Diskrepanzen zu Simulationsergebnissen festgestellt.

Anhang A

Fouriertransformation der Bildintensität

Der Übersichtlichkeit halber wird die Transformation in einer Dimension durchgeführt. Die Wellenfunktion heie $\Psi(r)$, deren Transformierte sei $F(k)$:

$$\mathcal{F}(\Psi(r)) = F(k) = \int \Psi(r) \exp(2\pi i k r) dr \quad (1.1)$$

$$\mathcal{F}^{-1}(F(k)) = \Psi(r) = \int F(k) \exp(-2\pi i k r) dk \quad (1.2)$$

Die Fouriertransformierte der Bildintensit $\Psi\Psi^*$ lautet damit:

$$\begin{aligned} \mathcal{F}(\Psi\Psi^*) = I(k) &= \int \Psi(r) \Psi^*(r) \exp(2\pi i k r) dr \\ &= \int \Psi(r) \underbrace{\left[\int F^*(k') \exp(2\pi i k' r) dk' \right]}_{\Psi^*(r)} \exp(2\pi i k r) dr \\ &= \int \underbrace{\left[\int \Psi(r) \exp[2\pi i (k + k') r] dr \right]}_{F(k+k')} F^*(k') dk' \\ &= \int F(k + k') F^*(k') dk' \end{aligned} \quad (1.3)$$

Diese auf den ersten Blick nicht leicht einsichtige Autokorrelation lt sich besser an einem diskreten Fourierspektrum erklren:

$$I(k) = \sum_{k'} F(k + k') F^*(k') \quad (1.4)$$

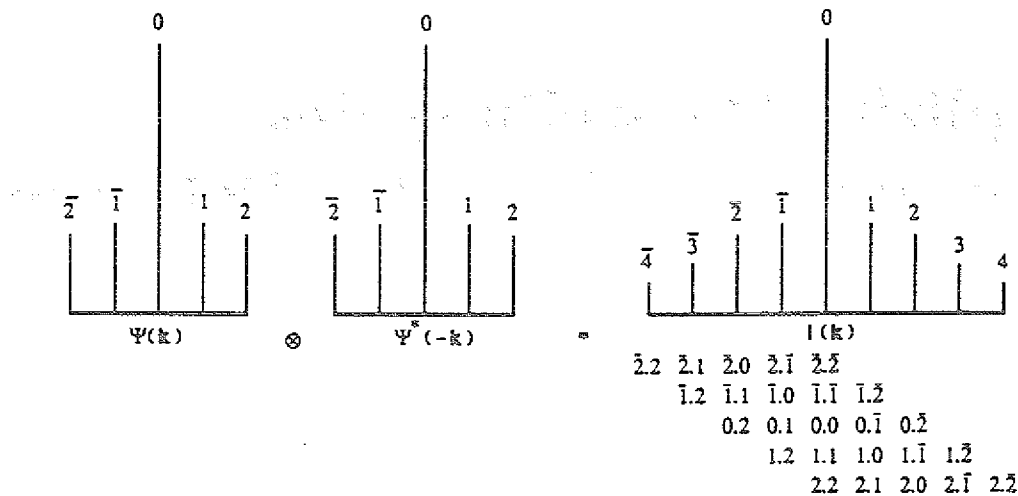
Der Summand ist hier der Interferenzbeitrag zweier Strahlen bzw. Fourierwellen mit:

$$\begin{aligned} \Psi_{k+k'} &= F(k + k') \exp[2\pi i (k + k') r] \\ \Psi_{k'}^* &= F^*(k') \exp(-2\pi i k' r) \end{aligned} \quad (1.5)$$

Bei der Betragsquadratbildung $\Psi\Psi^*$ treten Interferenzterme der folgenden Form auf:

$$\Psi_{k+k'} \Psi_{k'}^* = F(k + k') F^*(k') \exp(2\pi i k r) \quad (1.6)$$

In der *Bildintensität* bewirkt ein solcher Interferenzterm das Auftreten einer Raumfrequenz k . Im allgemeinen gibt es viele Möglichkeiten zwei Strahlen zu finden, deren Differenz der Wellenvektoren gerade eine bestimmte Bildfrequenz k ergibt. Die Autokorrelation summiert über alle möglichen Strahlenpaare, deren Interferenzbeitrag zur Bildintensität gerade in einer bestimmten Raumfrequenz k resultiert.



In der obigen Abbildung ist anschaulich dargestellt, wie die Autokorrelation einer aus den fünf Strahlen $2, 1, 0, 1, 2$ bestehenden Wellenfunktion Ψ zu einem aus neun Fourierkomponenten bestehenden Bildfrequenzspektrum führt. Die Interferenzbeiträge zweier Strahlen x und y sind hierbei unter jeder Bildfrequenz in der Notation $x.y$ aufgelistet. In einer Dimension führt die Autokorrelation einer Wellenfunktion aus N Strahlen zu $2N - 1$ Fourierkomponenten des Bildes, wobei insgesamt N^2 Interferenzterme auftreten. Der im folgenden dargestellte Transmissions-Kreuzkoeffizient ist eine komplexe Zahl, die jeden dieser Interferenzterme unterschiedlich „gewichtet“.

Anhang B

Explizite Darstellung des Transmissions-Kreuzkoeffizienten

$$T(\vec{k}'', \vec{k}', Z) = a(\vec{k}'') a(\vec{k}') \exp \left\{ -2\pi i \left[\chi(\vec{k}'', Z) - \chi(\vec{k}', Z) \right] \right\} \\ \times P(\vec{k}'', \vec{k}', Z) E_s(\vec{k}'', \vec{k}', Z) E_f(\vec{k}'', \vec{k}') E_x(\vec{k}'', \vec{k}') \quad (2.1)$$

mit den Funktionen:

$$P(\vec{k}'', \vec{k}', Z) = \exp \left[-2\pi i (\pi q_0 \Delta)^2 (\chi''_Z - \chi'_Z) (\chi''_k - \chi'_k, \chi''_{kZ} - \chi'_{kZ}) / u \right] \quad (2.2)$$

$$E_s(\vec{k}'', \vec{k}', Z) = \exp \left[-(\pi q_0)^2 (\chi''_k - \chi'_k)^2 / u \right] \quad (2.3)$$

$$E_f(\vec{k}'', \vec{k}') = \exp \left[-(\pi \Delta)^2 (\chi''_Z - \chi'_Z)^2 / u \right] \quad (2.4)$$

$$E_x(\vec{k}'', \vec{k}') = \frac{1}{\sqrt{u}} \exp \left[-(\pi^2 q_0^2 \Delta)^2 \{ (\chi''_k - \chi'_k) \times (\chi''_{kZ} - \chi'_{kZ}) \}^2 / u \right] \quad (2.5)$$

wobei $\chi(\vec{k}', Z)$, χ'_k , χ'_Z , χ'_{kZ} und u folgendermaßen definiert sind:

$$\chi(\vec{k}', Z) = \frac{1}{4} C_s \lambda^3 |\vec{k}'|^4 + \frac{1}{2} Z \lambda |\vec{k}'|^2 \quad (2.6)$$

$$\chi'_k = \left. \frac{\partial \chi(\vec{k}, Z)}{\partial \vec{k}} \right|_{\vec{k}=\vec{k}', Z=Z'} = (C_s \lambda^2 |\vec{k}'|^2 + Z') \lambda \vec{k}' \quad (2.7)$$

$$\chi'_Z = \left. \frac{\partial \chi(\vec{k}, Z)}{\partial Z} \right|_{\vec{k}=\vec{k}', Z=Z'} = \frac{\lambda}{2} |\vec{k}'|^2 \quad (2.8)$$

$$\chi'_{kZ} = \left. \frac{\partial^2 \chi(\vec{k}, Z)}{\partial \vec{k} \partial Z} \right|_{\vec{k}=\vec{k}', Z=Z'} = \lambda \vec{k}' \quad (2.9)$$

$$u = 1 + (\pi q_0 \Delta)^2 (\chi''_{kZ} - \chi'_{kZ})^2 \\ = 1 + \alpha (\vec{k}'' - \vec{k}')^2 \quad \text{mit} \quad \alpha = (\pi \lambda q_0 \Delta)^2 \quad (2.10)$$

Literaturverzeichnis

- [1] D. Van Dyck and A. F. de Jong. Ultimate resolution and information in electron microscopy: general principles. *Ultramicroscopy*, 47:266–281, (1992).
- [2] John C. Barry. Semiquantitative image matching in HRTEM. in: *Computer Simulation of Electron Microscope Diffraction and Images*, Eds. W. Krakow and M. A. O'Keefe (The Minerals, Metals and Materials Society, Warrendale), (1989).
- [3] A. Thust and K. Urban. Quantitative high-speed matching of high-resolution electron microscopy images. *Ultramicroscopy*, 45:23–42, (1992).
- [4] Wayne E. King and Geoffrey H. Campbell. Determination of thickness and defocus by quantitative comparison of experimental and simulated high-resolution images. *Ultramicroscopy*, 51:128–135, (1993).
- [5] Max T. Otten and Wim M. J. Coene. High-resolution imaging on a field emission TEM. *Ultramicroscopy*, 48:77–91, (1993).
- [6] S. Kujawa and D. Krah. Performance of a low-noise CCD camera adapted to a transmission electron microscope. *Ultramicroscopy*, 46:395–403, (1992).
- [7] Hannes Lichte. Electron holography approaching atomic resolution. *Ultramicroscopy*, 20:293–304, (1986).
- [8] Edgar Völkl and Hannes Lichte. Electron holograms for subångström point resolution. *Ultramicroscopy*, 32:177–180, (1990).
- [9] D. Van Dyck and M. Op de Beeck. New direct methods for phase and structure retrieval in HREM. In *Proc. 12th Int. Congr. for Electron Microscopy*, (1990).
- [10] W. Coene, G. Janssen, M. Op de Beeck, and D. Van Dyck. Phase retrieval through focus variation for ultra-resolution in field-emission transmission electron microscopy. *Phys. Rev. Letters*, 69:3743–3746, (1992).
- [11] D. Van Dyck, M. Op de Beeck, and W. Coene. A new approach to object wavefunction reconstruction in electron microscopy. *Optik*, 93 No. 3:103–107, (1993).

- [12] E. Ruedl, P. Delavignette, and S. Amelinckx. Electron diffraction and electron microscopic study of long- and short-range order in Ni_4Mo and of the substructure resulting from ordering. *phys. stat. sol.*, **28**:305–328, (1968).
- [13] S. Banerjee, K. Urban, and M. Wilkens. Order-disorder transformation in Ni_4Mo under electron irradiation in a high voltage electron microscope. *Acta metall.*, **32**:299–311, (1984).
- [14] J. Mayer and K. Urban. Spinodal ordering in Ni_4Mo . *Acta metall.*, **33**:539–543, (1985).
- [15] J. Mayer. Untersuchung der spinodalen Ordnungseinstellung in Ni_4Mo mit Hilfe der Hochspannungselektronenmikroskopie. *Diplomarbeit Universität Stuttgart*, (1984).
- [16] G. van Tendeloo and S. Amelinckx. Lattice relaxation at glide antiphase boundaries and stacking faults in Ni_3Mo and Ni_4Mo . *phys. stat. sol. (a)*, **27**:93–98, (1975).
- [17] K. H. Lee, K. Hiraga, D. Shindo, and M. Hirabayashi. High resolution electron microscopic study of the ordering processes in Ni_4Mo alloy. *Acta metall.*, **36**:641–649, (1988).
- [18] P. Schwander. Nahordnung und Versetzungen in einer Ni-Mo Legierung. *Dissertation Nr. 9227, ETH Zürich*, (1990).
- [19] P. Villars and L. D. Calvert. *Pearson's Handbook of Crystallographic Data for Intermetallic Phases*. American Society for Metals, 1986.
- [20] *International Tables for Crystallography*, volume A. D. Reidel Publishing Company, Dordrecht, (1987).
- [21] *International Tables for X-Ray Crystallography*, volume III. D. Reidel Publishing Company, Dordrecht, (1985).
- [22] M. A. Beno, L. Soderholm, D. W. Capone II, D. G. Hinks, J. D. Jorgensen, J. D. Grace, I. K. Schuller, C. U. Segre, and K. Zhang. Structure of the single-phase high-temperature superconductor $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$. *Appl. Phys. Lett.*, **51**:57–59, (1987).
- [23] K. Yvon and M. Francois. Crystal structures of high- T_c oxides. *Z. Phys. B-Condensed Matter*, **76**:413–444, (1989).
- [24] M. Francois, A. Junod, K. Yvon, A. W. Hewat, J. J. Capponi, P. Strobel, M. Marezio, and P. Fischer. A study of the CuO chains in the high T_c superconductor $\text{YBa}_2\text{Cu}_3\text{O}_7$ by high resolution neutron powder diffraction. *Solid State Commun.*, **66**:1117–1125, (1988).

- [25] J. D. Jorgensen, D. G. Hinks, P. G. Radaelli, S. Pei, P. Lightfoot, B. Dabrowski, C. U. Segre, and B. A. Hunter. Defects, defect ordering, structural coherence and superconductivity in the 123 copper oxides. *Physica C*, **185**:184–189, (1991).
- [26] C. L. Jia, B. Kabius, H. Soltner, U. Poppe, K. Urban, J. Schubert, and Ch. Buchal. A study of the microstructure of epitaxial heterostructures of $\text{YBa}_2\text{Cu}_3\text{O}_7$ and $\text{PrBa}_2\text{Cu}_3\text{O}_7$. *Physica C*, **167**:463–471, (1990).
- [27] C. L. Jia, B. Kabius, K. Urban, K. Herrman, G. J. Cui, J. Schubert, W. Zander, A. I. Braginski, and C. Heiden. Microstructure of epitaxial $\text{YBa}_2\text{Cu}_3\text{O}_7$ films on step-edge SrTiO_3 substrates. *Physica C*, **175**:545–554, (1991).
- [28] C. L. Jia. Investigation of thin films, heterostructures and devices of ceramic superconductors by means of high-resolution electron microscopy. *Berichte des Forschungszentrum Jülich*, Nr. 2809, (1993).
- [29] R. Ramesh, D. M. Hwang, T. S. Ravi, A. Inam, X. D. Wu, and T. Venkatesan. Polytypoidic superlattices in Y–Ba–Cu–O thin films. *Physica C*, **171**:14–18, (1990).
- [30] R. Ramesh, D. M. Hwang, J. B. Barner, L. Nazar, T. S. Ravi, A. Inam, B. Dutta, X. D. Wu, and T. Venkatesan. Defect structure of laser deposited Y–Ba–Cu–O thin films on single crystal MgO substrate. *J. Mater. Res.*, **5**:704–716, (1990).
- [31] Y. Gao, K. L. Merkle, G. Bai, H. L. M. Chang, and D. J. Lam. [001] tilt grain boundaries in $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ thin films. *Ultramicroscopy*, **37**:326–340, (1991).
- [32] Y. Gao, G. Bai, D. J. Lam, and K. L. Merkle. Microstructure and defects in a-axis oriented $\text{YBa}_2\text{Cu}_3\text{O}_{7-x}$ thin films. *Physica C*, **173**:487–500, (1991).
- [33] W. Krakow and T. M. Shaw. Evaluation of high-resolution electron microscopy as a method for studying Y–Ba–Cu–O superconductors. *Journal of electron microscopy technique*, **8**:273–284, (1988).
- [34] H. W. Zandbergen, R. Gronsky, K. Wang, and G. Thomas. Simulation and processing of high resolution images of defects and interfaces in superconducting $\text{YBa}_2\text{Cu}_3\text{O}_7$. *Mat. Res. Soc. Symp. Proc.*, **99**:553–556, (1988).
- [35] A. Ourmazd and J. C. H. Spence. Detection of oxygen ordering in superconducting cuprates. *Nature*, **329**:425–427, (1987).
- [36] A. Ourmazd, J. C. H. Spence, J. M. Zuo, and C. H. Li. Structural characterization of $\text{Ba}_2\text{YCu}_3\text{O}_7$ by high resolution transmission electron microscopy. *Journal of electron microscopy technique*, **8**:251–262, (1988).
- [37] N. P. Huxford, D. J. Eaglesham, and C. J. Humphreys. Limits on quantitative information from high-resolution electron microscopy of $\text{YBa}_2\text{Cu}_3\text{O}_7$ superconductors. *Nature*, **329** No. 6142:812–813, (1987).

- [38] C. L. Jia, A. Thust, G. Jakob, and K. Urban. Chemically sensitive imaging of $\text{YBa}_2\text{Cu}_3\text{O}_7/\text{PrBa}_2\text{Cu}_3\text{O}_7$ superlattices by means of high-resolution electron microscopy. *Ultramicroscopy*, 49:330–343, (1993).
- [39] L. Reimer. *Transmission Electron Microscopy*. Springer Verlag, 1984.
- [40] P. A. Doyle and P. S. Turner. Relativistic Hartree-Fock X-ray and electron scattering factors. *Acta Cryst.*, A 24:390–397, (1968).
- [41] G. H. Smith and R. E. Burge. *Acta Cryst.*, 15:182, (1962).
- [42] H. Bethe. *Ann. Phys.*, 87:55, (1928).
- [43] C. J. Humphreys. The scattering of fast electrons by crystals. *Rep. Prog. Phys.*, 42:1864, (1979).
- [44] J. M. Cowley and A. F. Moodie. The scattering of electrons by atoms and crystals. I. A new theoretical approach. *Acta Cryst.*, 10:609 – 619, (1957).
- [45] L. C. Qin and K. Urban. Electron diffraction and lattice image simulations with the inclusion of HOLZ reflections. *Ultramicroscopy*, 33:159 – 166, (1990).
- [46] P. A. Stadelmann. EMS - a software package for electron diffraction analysis and HREM image simulation in materials science. *Ultramicroscopy*, 21:131–145, (1987).
- [47] K. Ishizuka and N. Uyeda. A new theoretical and practical approach to the multislice method. *Acta Cryst.*, A 33:740 – 749, (1977).
- [48] P. Goodman and A. F. Moodie. Numerical evaluation of n-beam wave functions in electron scattering by the multislice method. *Acta Cryst.*, A 30:280–290, (1974).
- [49] O. Scherzer. The theoretical resolution limit of the electron microscope. *J. Appl. Physics*, 20:20, (1949).
- [50] K. Heinemann. In-situ measurement of objective lens data of a high-resolution electron microscope. *Optik*, 34:113–128, (1971).
- [51] W. O. Saxton. Spatial coherence in axial high resolution conventional electron microscopy. *Optik*, 49:51–62, (1977).
- [52] M. Troyon. A method for determining the illumination divergence from electron micrographs. *Optik*, 49:247–251, (1977).
- [53] J. Frank. The envelope of electron microscopic transfer functions for partially coherent illumination. *Optik*, 38:519–536, (1973).
- [54] R. H. Wade and J. Frank. Electron microscopic transfer functions for partially coherent illumination and chromatic defocus spread. *Optik*, 49:81–92, (1977).

- [55] K. Ishizuka. Contrast transfer of crystal images in TEM. *Ultramicroscopy*, 5:55-65, (1979).
- [56] M. Troyon. Use of a field emission gun in transmission electron microscopy. *Optik*, 46:439-450, (1976).
- [57] J. C. H. Spence. *Experimental High Resolution Electron Microscopy*. Clarendon Press, Oxford, 1981.
- [58] J. Frank. Determination of source size and energy spread from electron micrographs using the method of Young's fringes. *Optik*, 44:379-391, (1975).
- [59] A. Bourret and J. M. Penisson. Atomic imaging with the JEOL JEM-4000EX electron microscope. *JEOL news*, 25E:2-7, (1987).
- [60] J. E. Bonevich and L. D. Marks. Contrast transfer theory for non-linear imaging. *Ultramicroscopy*, 26:313-320, (1988).
- [61] J. E. Bonevich and L. D. Marks. Contrast transfer theory for non-linear imaging (addendum). *Ultramicroscopy*, 32:297, (1990).
- [62] F. Zemlin, K. Weiss, P. Schiske, W. Kunath, and K. H. Herrmann. Coma-free alignment of high resolution electron microscopes with the aid of optical diffractograms. *Ultramicroscopy*, 3:49-60, (1978).
- [63] F. Zemlin. A practical procedure for alignment of a high resolution electron microscope. *Ultramicroscopy*, 4:241-245, (1979).
- [64] G. Möbus and M. Rühle. A new procedure for the determination of the chromatic contrast transfer envelope of electron microscopes. *Optik*, 93 No.3:108-118, (1993).
- [65] F. Thon. *Z. Naturforsch.*, 21a:476, (1966).
- [66] P. M. Kelly, A. Jostsons, and J. G. Napier. The determination of foil thickness by scanning transmission electron microscopy. *phys. stat. sol. (a)*, 31:771-780, (1975).
- [67] C. J. D. Hetherington, M. A. O'Keefe, and R. Kilaas. Simulation of high resolution images of wedge-shaped crystals. in: *Computer Simulation of Electron Microscope Diffraction and Images*, Eds. W. Krakow and M. A. O'Keefe (The Minerals, Metals and Materials Society, Warrendale), (1989).
- [68] I. Daberkow, K. H. Herrmann, L. Liu, and W. D. Rau. Performance of electron image converters with YAG single crystal screen and CCD sensor. *Ultramicroscopy*, 38:215-223, (1991).
- [69] O. L. Krivanek and P. E. Mooney. Applications of slow-scan CCD cameras in transmission electron microscopy. *Ultramicroscopy*, 49:95-108, (1993).

- [70] J. S. Lim. *Two-Dimensional Signal and Image Processing*. Prentice-Hall, Englewood Cliffs, NJ, (1990).
- [71] Marin van Heel, Michael Schatz, and Elena Orlova. Correlation functions revisited. *Ultramicroscopy*, 46:307–316, (1992).
- [72] A. Ourmazd, F. H. Baumann, M. Bode, and Y. Kim. Quantitative chemical lattice imaging: theory and practice. *Ultramicroscopy*, 34:237–255, (1990).
- [73] S. Thoma and H. Cerva. New methods for qualitative and quantitative analysis of the GaAs/AlGaAs interface by high-resolution electron microscopy. *Ultramicroscopy*, 38:265–289, (1991).
- [74] A. F. De Jong, W. Coene, and D. Van Dyck. Image processing of HRTEM images with non-periodic features. *Ultramicroscopy*, 27:53–66, (1989).
- [75] P. Pradère, J. F. Revol, L. Nguyen, and R. St. John Manley. Lattice imaging of poly-4-methyl-pentene-1 single crystals: Use and misuse of fourier averaging techniques. *Ultramicroscopy*, 25:69–80, (1988).
- [76] P. Pradère and E. L. Thomas. Image processing of partially periodic lattice images of polymers: The study of crystal defects. *Ultramicroscopy*, 32:149–168, (1990).
- [77] A. Ourmazd, D. W. Taylor, M. Bode, and Y. Kim. Quantifying the information content of lattice images. *Science*, 246:1571–1577, (1989).
- [78] G. Möbus, G. Necker, and M. Rühle. Adaptive fourier-filtering technique for quantitative evaluation of high-resolution electron micrographs of interfaces. *Ultramicroscopy*, 49:46–65, (1993).
- [79] R. W. Ramirez. *The FFT – Fundamentals and Concepts*. Prentice-Hall, Englewood Cliffs, NJ, (1985).
- [80] J. W. Cooley and J. W. Tukey. *Mathematics of Computation*, 19:297, (1965).
- [81] R. Kilaas and R. Gronsky. *Ultramicroscopy*, 16:193, (1985).
- [82] S. Paciornik, R. Kilaas, and U. Dahmen. Assessment of specimen noise in HREM images of simple structures. *Ultramicroscopy*, 50:255–262, (1993).
- [83] C. E. Shannon. *Bell System Tech. J.*, 27:379, (1948).
- [84] U. Aebi, P. R. Smith, J. Dubochet, C. Henry, and E. Kellenberger. A study of the structure of the T-layer of bacillus brevis. *Journal of Supramolecular Structure*, 1:498–522, (1973).
- [85] W. O. Saxton, R. Dürr, and W. Baumeister. From lattice distortion to molecular distortion: characterising and exploiting crystal deformation. *Ultramicroscopy*, 46:287–306, (1992).

- [86] S. Hovmöller. Structure analysis by crystallographic image processing. *Microsc. Microanal. Microstruct.*, 1:423-431, (1990).
- [87] S. Hovmöller, A. Sjögren, G. Farrants, M. Sundberg, and B. O. Marinder. Accurate atomic positions from electron microscopy. *Nature*, 311:238-241, (1984).
- [88] D. X. Li and S. Hovmöller. The crystal structure of $\text{Na}_3\text{Nb}_{12}\text{O}_{31}\text{F}$ determined by HREM and image processing. *J. Solid State Chem.*, 73:5-10, (1988).
- [89] D. N. Wang, S. Hovmöller, L. Kihlberg, and M. Sundberg. Structure determination and correction for distortions in HREM by crystallographic image processing. *Ultramicroscopy*, 25:303-316, (1988).
- [90] L. Reimer and R. Rennekamp. Imaging and recording of multiple scattering effects by angular-resolved electron energy-loss spectroscopy. *Ultramicroscopy*, 28:258-265, (1989).
- [91] E. Zeitler. The photographic emulsion as analog recorder for electrons. *Ultramicroscopy*, 46:405-416, (1992).
- [92] G. Y. Fan and J. M. Cowley. Auto-correlation analysis of high resolution electron micrographs of near-amorphous thin films. *Ultramicroscopy*, 17:345-356, (1985).
- [93] J. C. H. Spence. *Experimental High Resolution Electron Microscopy*, page 139. Oxford University Press, New York, 1988.
- [94] D. J. Smith, W. O. Saxton, M. A. O'Keefe, G. J. Wood, and W. M. Stobbs. *Ultramicroscopy*, 11:263, (1983).
- [95] R. Henderson, J. M. Baldwin, K. H. Downing, J. Lepault, and F. Zemlin. Structure of purple membrane from halobacterium halobium: Recording, measurement and evaluation of electron micrographs at 3.5 Å resolution. *Ultramicroscopy*, 19:147-178, (1986).
- [96] E. H. Linfoot. *Qualitätsbewertung optischer Bilder*. Vieweg Braunschweig, (1960).
- [97] D. Van Dyck, J. Danckaert, W. Coene, E. Selderslaghs, D. Broddin, J. Van Landuyt, and S. Amelinckx. The atom column approximation in dynamical electron diffraction calculations. in: *Computer Simulation of Electron Microscope Diffraction and Images*, Eds. W. Krakow and M. A. O'Keefe (The Minerals, Metals and Materials Society, Warrendale), (1989).
- [98] D. Van Dyck. A simple real-space channeling theory for electron diffraction and HREM. In *Proc. 12th Int. Congr. for Electron Microscopy*, (1990).
- [99] D. Van Dyck and W. Coene. Direct structure retrieval from high resolution electron micrographs. *Scanning Microscopy Supplement*, 2:131-137, (1988).

- [100] D. Gabor. *Proc. Roy. Soc., Ser. A* 197:454, (1949).
- [101] Hannes Lichte. Optimum focus for taking electron holograms. *Ultramicroscopy*, 38:13–22, (1991).
- [102] Hannes Lichte. Electron holography improves the wave-optical facilities in electron microscopy. *Philips Electron Optics Bulletin*, 130:29–35, (1992).
- [103] Hannes Lichte. Parameters for high-resolution electron holography. *Ultramicroscopy*, 51:15–20, (1993).
- [104] David C. Joy, Y. S. Zhang, X. Zhang, T. Hashimoto, R. D. Bunn, L. Allard, and T. A. Nolan. Practical aspects of electron holography. *Ultramicroscopy*, 51:1–14, (1993).
- [105] M. A. Gribelyuk and J. M. Cowley. Imaging conditions for off-axis transmission electron holography. *Ultramicroscopy*, 50:29–40, (1993).
- [106] P. Schiske. in: *Image Processing and Computer Aided Design in Electron Optics*, (Academic Press, New York), (1973).
- [107] Y. Liu, S. Xiang, H. Fan, T. Tang, F. Li, Q. Pan, N. Uyeda, and Y. Fujiyoshi. Image deconvolution of a single high-resolution electron micrograph. *Acta Cryst.*, A46:459–463, (1990).
- [108] J. J. Hu and F. H. Li. Maximum entropy image deconvolution in high resolution electron microscopy. *Ultramicroscopy*, 35:339–350, (1991).
- [109] A. J. J. Drenth, A. M. Huizer, and H. A. Ferwerda. The problem of phase retrieval in light and electron microscopy of strong objects. *Optica Acta*, 22:615–628, (1975).
- [110] D. L. Misell. An examination of an iterative method for the solution of the phase problem in optics and electron optics. *J. Phys. D: Appl. Phys.*, 6:2200–2216, (1973).
- [111] R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–247, (1972).
- [112] A. M. J. Huizer, A. J. J. Drenth, and H. A. Ferwerda. On phase retrieval in electron microscopy from image and diffraction pattern. *Optik*, 45:303–316, (1976).
- [113] E. J. Kirkland. Improved high resolution image processing of bright field electron micrographs (theory). *Ultramicroscopy*, 15:151–172, (1984).
- [114] E. J. Kirkland, B. M. Siegel, N. Uyeda, and Y. Fujiyoshi. Improved high resolution image processing of bright field electron micrographs (experiment). *Ultramicroscopy*, 17:87–104, (1985).

- [115] E. J. Kirkland. High resolution image processing of electron micrographs. *Scanning Microscopy Supplement*, 2:139–147, (1988).
- [116] M. A. Gribelyuk and J. L. Hutchison. On the iterative restoration of the exit plane wave function from defocus series in HREM. *Ultramicroscopy*, 45:127–143, (1992).
- [117] L. Davis. *Genetic Algorithms and Simulated Annealing*. Pitman, London, (1987).
- [118] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equations of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087–1091, (1953).
- [119] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, (1983).
- [120] S. Kirkpatrick. Optimization by simulated annealing: Quantitative studies. *Journal of Statistical Physics*, 34:975–986, (1984).
- [121] J. H. Holland. *Adaption in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, (1975).
- [122] W. M. J. Coene and T. J. J. Denteneer. Improved methods for the determination of the spherical aberration coefficient in high-resolution electron microscopy from micrographs of an amorphous object. *Ultramicroscopy*, 38:225–233, (1991).
- [123] A. J. Koster and A. F. de Jong. Measurement of the spherical aberration coefficient of transmission electron microscopes by beam-tilt-induced image displacements. *Ultramicroscopy*, 38:235–240, (1991).

It is essential for the company to ensure that all financial data is recorded correctly and that the accounting system is properly maintained. This will help to ensure that the company's financial statements are accurate and reliable.

It is essential for the company to ensure that all financial data is recorded correctly and that the accounting system is properly maintained. This will help to ensure that the company's financial statements are accurate and reliable.

It is essential for the company to ensure that all financial data is recorded correctly and that the accounting system is properly maintained. This will help to ensure that the company's financial statements are accurate and reliable.

It is essential for the company to ensure that all financial data is recorded correctly and that the accounting system is properly maintained. This will help to ensure that the company's financial statements are accurate and reliable.

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, von denen ich während der Anfertigung dieser Arbeit Hilfe und Unterstützung erfahren habe.

Bei Herrn Prof. Dr. Urban bedanke ich mich für die Möglichkeit, diese Arbeit an seinem Institut durchführen zu können, sowie für seine erfahrene und fachkundige Betreuung, durch die er mich stets motiviert und unterstützt hat.

Bei Herrn Dipl.-Phys. M. Lentzen möchte ich mich in besonderem Maße für die äußerst produktive und herzerfrischende Zusammenarbeit bei der gemeinsamen Entwicklung der Rekonstruktionsalgorithmen bedanken.

Herrn Dr. B. Kabius danke ich für seinen kompetenten Rat und seine stete Diskussionsbereitschaft bei vielfältigen Fragen bezüglich Abbildung und Bildsimulation.

Frau Dipl.-Phys. M. Seo danke ich für ihre Einweisung in die Geheimnisse der Präparation von Querschnittsproben von Hochtemperatursupraleitern.

Herrn Dr. C. L. Jia danke ich für die vielfältigen Hinweise und Tips zur Probenpräparation, die mir von großem Nutzen waren. Ebenso möchte ich mich bei ihm für den bereitwilligen Erfahrungsaustausch bezüglich Abbildung und Simulation auf dem Gebiet der Hochtemperatursupraleiter bedanken.

Herrn Dipl.-Phys. M. Wollgarten danke ich für die Überlassung der von ihm geschriebenen Graphikprogramme, sowie für seine stete Hilfsbereitschaft bei tückischen Problemen auf den Gebieten von Graphik und Textverarbeitung.

Bei Frau Dr. M. Luysberg sowie den Herrn Dipl.-Phys. Th. Walther und Dr. D. Stenkamp bedanke ich mich für zahlreiche Diskussionen und Anregungen bezüglich Bildsimulation und Bildverarbeitung.

Frau W. Sybertz und Frau C. Scholz danke ich für ihren Rat und ihre Hilfsbereitschaft auf dem Gebiet der Probenpräparation.

Bei Frau G. Waßenhoven und Frau R. Schieren bedanke ich mich für die sorgfältige und prompte Erledigung der anfallenden Photoarbeiten.

Herrn Dipl.-Ing. J. Hanßen danke ich für seine zuvorkommende und kompetente Hilfe bei technischen Problemen rund um Labor und Mikroskop.

Neben dem fachbezogenen Dank möchte ich den derzeitigen und ehemaligen Doktoranden, Diplomanden und Gästen des Instituts für Mikrostrukturforschung sowie den fest angestellten Mitarbeitern für die menschlich angenehme Atmosphäre, die mir viel Freude bereitet hat, danken.

Introduction

The purpose of this document is to provide a comprehensive overview of the project's objectives, scope, and deliverables. It serves as a reference for all stakeholders involved in the project.

The project aims to develop a new software application that will streamline the workflow of the department. The application will be designed to be user-friendly and efficient, ensuring that all team members can utilize it effectively.

The project is divided into several phases, including planning, development, testing, and deployment. Each phase has specific tasks and milestones that must be completed to ensure the project's success.

The project team consists of several members, each with specific responsibilities. The team will work closely together to ensure that all project goals are met and that the application is delivered on time.

The project budget is estimated to be \$100,000. This budget includes all costs associated with the project, including personnel, materials, and other resources.

The project timeline is as follows: Planning (2 weeks), Development (8 weeks), Testing (4 weeks), and Deployment (2 weeks). The project is expected to be completed by the end of the year.

The project will be managed using a project management tool. This tool will allow the team to track progress, manage resources, and communicate effectively throughout the project.

The project will be reviewed regularly to ensure that it is on track and that any issues are identified and resolved promptly. The project manager will provide regular updates to the stakeholders.

The project is expected to have a positive impact on the department's workflow and efficiency. It will also provide valuable experience for the team members involved in the project.

The project is a high-priority initiative for the department. It is essential that all team members are committed to the project and work together to ensure its success.

The project is a complex task that requires careful planning and execution. It is important to stay organized and focused throughout the project to ensure that all goals are met.

The project is a challenging but rewarding task. It provides an opportunity for the team to learn, grow, and achieve something great together. We are confident that the project will be a success.

Jül-2978
Oktober 1994
ISSN 0944-2952